

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 April 2002 (25.04.2002)

PCT

(10) International Publication Number  
**WO 02/32943 A2**

(51) International Patent Classification<sup>7</sup>: **C07K 14/16**

(21) International Application Number: **PCT/US01/25721**

(22) International Filing Date: 14 August 2001 (14.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/225,097 14 August 2000 (14.08.2000) US  
60/252,115 14 November 2000 (14.11.2000) US  
60/279,257 28 March 2001 (28.03.2001) US

(71) Applicants (for all designated States except US): **THE GOVERNMENT OF THE UNITED STATES OF AMERICA**, as represented by **THE SECRETARY**,

**DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]**; National Institutes of Health, Office of Technology Transfer, 6011 Executive Blvd., Suite 325, Rockville, MD 20852-3804 (US). **CHADRABARTI, Bimal, K.** [IN/US]; 1147 McIntyre Drive, Ann Arbor, MI 48105 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **NABEL, Gary, J.** [US/US]; 2520 30th Street, NW, Washington, DC 20008 (US). **HUANG, Yue** [CA/US]; 413 West Side Drive, #302, Gaithersburg, MD 20878 (US).

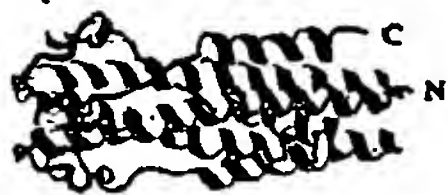
(74) Agent: **ALTMAN, Daniel, E.**; Knobbe, Martens, Olson and Bear, LLP, 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA,

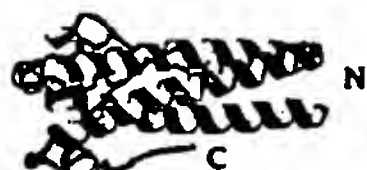
[Continued on next page]

(54) Title: **MODIFICATIONS OF HIV ENV, GAG, AND POL ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION**

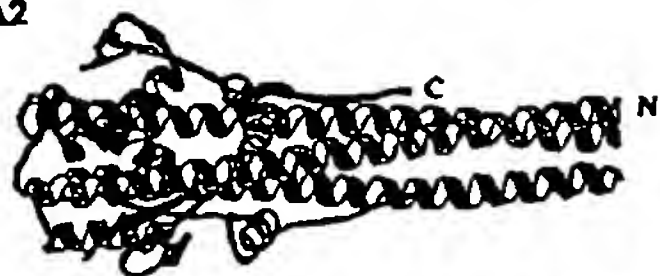
A Ebola GP2



B MoMuLv Mo-55



C low-pH-treated influenza HA2



D protease resistant core of HIV gp41



E SIV gp41 (NMR)



F SNAREs



(57) Abstract: Modified HIV Env, Gag, Pol, or Nef DNA with improved ability to elicit antibody and CTL responses to HIV antigens have been identified as prototype immunogens for the treatment and prevention of HIV infections.

WO 02/32943 A2



CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EC, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report*

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**Modifications of HIV Env, Gag, and Pol Enhance Immunogenicity  
for Genetic Immunization**

5

**Field of the Invention**

The present invention relates to the field of molecular biology. The present invention discloses modified HIV Env, Gag, Pol, and Nef proteins, related nucleotide sequences, and usage for genetic immunization.

**Background of the Invention**

10

Protective immunity against human immunodeficiency virus-1 (HIV-1) is likely to require recognition of linear and conformation epitopes from multiple HIV antigens. Whether these responses can be elicited more effectively by virion-like structures or fused CTL epitopes is unknown.

15

The immune response to HIV infection in long-term non-progressors (Cao Y et al., 1995, N Engl J Med, 332:201-8; Pantaleo G et al., 1995, N Engl J Med, 332:209-16) and HIV-exposed sex workers suggests that specific viral immunity may limit infection and the symptoms of disease. No single characteristic yet correlates with protective immunity, but studies in non-human of primates suggest that both humoral and cellular immunity are required for this response. Depletion of cytotoxic T cells (CTLs) in chronically-infected macaques enhances viremia (Jin X et al., 1999, J Exp Med, 189:991-8; Schmitz JE et al., 1999, Science, 283:857-60). In humans, higher CTL responses correlate with lower viral load and stabilization of clinical symptoms (Musey L et al., 1997, N Engl J Med, 337:1267-74; Ogg GS et al., 1998, Science, 279:2103-6). In animal models, passive transfer of neutralizing antibodies can also contribute to protection against virus challenge (Burton DR, Montefiori DC, 1997, AIDS, 11 Suppl A:S87-S98; Gauduin MC et al., 1997, Nat Med, 3:1389-93; Moore JP, Ho DD., 1995, AIDS, 9 Suppl A:S117-S136; Muster T et al., 1994, J Virol, 68:4031-4; Muster T et al., 1993, J Virol, 67:6642-7; Poignard P et al., 1996, Immunol Today, 17:239-46; Sattentau QJ, 1996, Curr Opin Immunol, 8:540-5; Shibata R et al., 1999, Nat Med, 5:204-10). Neutralizing antibody responses can also be developed in HIV-infected individuals (Burton DR, Montefiori DC., 1997, AIDS, 11 Suppl A:S87-S98; Poignard P et al., 1996, Immunol Today, 17:239-46; Sattentau QJ., 1996, Curr Opin Immunol, 8:540-5) and are associated with lower viral loads in long-term non-progressors (Montefiori DC et al., 1996, J Infect Dis, 173:60-7). Though this neutralizing antibody response is

35

uncommon, it is directed largely against the Env protein of the virus (Robey WG et al., 1986, Proc Natl Acad Sci U S A, 83:7023-7; Steimer KS et al., 1991, Science, 254:105-8).

5 In early human vaccine trials, gp120 protein immunogens have yielded disappointing results: vaccine-induced antibodies have not been broadly neutralizing and have sometimes enhanced infection in vitro (Bolognesi DP, Matthews TJ, 1998, Nature, 391:638-9; Connor RI et al., 1998, J Virol, 72:1552-76; Haynes BF, 1996, Lancet, 348:933-7; VanCott TC et al., 1995, J Immunol, 155:4100-10). Monomeric gp120 loses oligomer-dependent epitopes and does not include sequences in the  
10 ectodomain of the gp41 that become exposed during virus entry (Broder CC et al., 1994, Proc Natl Acad Sci U S A, 91:11699-703; Moore JP, 1995, Nature, 376:115; Muster T et al., 1993, J Virol, 67:6642-7). It is assumed that broadly neutralizing antibodies bind to native gp120/gp41 complex on the surface of the virus rather than soluble gp120 (Burton DR, Montefiori DC, 1997, AIDS, 11 Suppl A:S87-S98).

15 The development of a cytotoxic T lymphocyte (CTL) response to viruses is often crucial to the outcome of infections. Lysis of infected cells prior to the production of progeny virions may limit virus burst size (Yang, O et al., 1996, J. Virol., 70:5799-5806), and HIV specific CD8<sup>+</sup> cytotoxic T lymphocytes (CTL) have been shown to be important in viral clearance and in the control of initial HIV-1  
20 spread (Borrow, P et al., 1994, J. Virol., 68:6103-6110; Yang, O et al., 1996, J. Virol. 70:5799-5806). CTL responses specific to HIV also contribute to reduction in viral load during acute and asymptomatic infection (Klein, MR, et al., 1995, J. Exp. Med. 181:1365-1372; Moss, PAH et al., 1995, Proc. Natl. Acad. Sci. USA, 92:5773-5777) and may be involved in protection against the establishment of persistent HIV  
25 infections (Rowland-Jones, SL et al., 1993, Lancet, 341:860-861; Rowland-Jones, SL et al., 1995, Nat. Med., 1:59-64). High-frequency CTL responses to HIV-1 correlated with low viral load and slow disease progression in chronically infected individuals (Musey, L et al., 1997, N. Engl. J. Med., 337:1267-1274; Ogg, GS et al., 1998, Science, 279:2103-2106.). More compelling evidence of an antiviral effect of  
30 CD8<sup>+</sup> cells was demonstrated in controlled studies in macaques, in which CD8<sup>+</sup> cells were depleted *in vivo* using a monoclonal antibody. The viral loads in these animals increased or decreased as the CD8<sup>+</sup> cells were depleted or reappeared, respectively (Jin, X et al., 1999, J. Exp. Med., 189:991-998; Schmitz, JE, et al., 1999, Science, 283:857-860). Therefore, induction of a CTL response specific to these proteins  
35 represents a desirable response in an HIV-1 vaccine.

HIV-1 internal structural and enzymatic proteins contain conserved domains that preserve their functions and thus exhibit less antigenic diversity that may elicit more effective CTL responses (Nixon, DF et al., 1988, *Nature*, **336**:484-487.). Efficient and durable CTL responses require endogenous antigen synthesis and processing. Current vaccine delivery techniques include immunization with live, attenuated viruses, inactivated recombinant virus infection (Letvin, NL, 1998, *Science*, **280**:1875-1880) or plasmid DNA expression vectors. A major obstacle in the induction of CTL responses with naked DNA or recombinant virus during development of an HIV vaccine is that the expression of HIV-1 structural and enzymatic genes is tightly regulated by the virus itself. The expression of these proteins is heavily dependent upon the existence of the Rev-responsive element (RRE) of HIV-1 in recombinant vectors (Cullen, BR, 1992, *Microbiol. Rev.*, **56**:375-394; Felber, BK et al., 1989, *Proc Natl Acad Sci U S A*, **86**:1495-1499). Poor expression is caused by the presence of AT rich inhibitory nucleotide sequences (INS) in the *gag*, *pol* and *env* genes, which inhibit the nuclear export and efficient expression of unspliced HIV1 mRNAs. Early studies of DNA vaccination against HIV in mice required the inclusion of Rev in their expression vectors (Lu, S et al., 1995, *Virology*, **209**:147-154; Okuda, K et al., 1995, *AIDS Res. Hum. Retroviruses*, **11**:933-943; Wang, B et al., 1993, *Proc. Natl. Acad. Sci. U. S. A* **90**:4156-4160), but modification of INS has been shown to facilitate Rev-independent expression of HIV-1 Gag (Qiu, J-T et al., 1999, *J. Virol.*, **73**:9145-9152; zur Megede, J et al., 2000, *J Virol* **74**:2628-2635), allowing detectable humoral and CTL responses against this protein (Qiu, J-T et al., 1999, *J. Virol.*, **73**:9145-9152). These modified HIV-1 Gag genes produced viral-like particles of the expected density and morphology and induced an immune response to HIV-1 Gag after DNA immunization in mice (zur Megede, J et al., 2000, *J Virol*, **74**:2628-2635).

#### Summary of the Invention

Protective immunity against human immunodeficiency virus-1 (HIV-1) is likely to require recognition of linear and conformation epitopes from multiple HIV antigens, and whether these responses can be elicited more effectively by virion-like structures or fused CTL epitopes was previously unknown. Herein is provided a modified HIV Env with deletions in the cleavage site, fusogenic domain, and spacing of heptad repeats 1 and 2 to expose the core protein for optimal antigen

presentation and recognition. Additionally we provide a Gag-Pol or Gag-Pol-Nef fusion protein that is a polyprotein designed to maximize epitope presentation. The invention extends the mutation in HIV Env, Gag, Pol, and Nef to any HIV clade or strain and to related proteins of other viruses. Different combinations, different  
5 orders, and different variations on a theme are envisioned, the theme being to optimize presentation of epitopes that generate broad CTL and antibody responses.

More particularly, we have investigated the effect of specific mutations in human immunodeficiency virus type 1 (HIV-1) envelope (Env) on humoral and cellular immune responses after DNA vaccination. Mice were injected with plasmid  
10 expression vectors encoding HIV Env with modifications of conserved glycosylation sites or different COOH-terminal mutations intended to mimic a fusion intermediate. Elimination of conserved glycosylation sites did not substantially enhance humoral or CTL immunity. In contrast, a modified gp140 with deletions in the cleavage site, fusogenic domain and spacing of heptad repeats 1 and 2 enhanced humoral  
15 immunity without reducing the efficacy of the CTL response. Because of its ability to stimulate the antibody response to native gp160 without affecting cellular immunity, this modified gp140 or a related derivative is envisioned to be a useful component of an AIDS vaccine.

In addition, we have examined the immune response to HIV-1 Gag and Pol  
20 after plasmid DNA immunization with Rev-independent expression vectors encoding various forms of these proteins. Immune responses were analyzed after vaccination with four expression vectors, including Gag alone or Gag-Pol, both of which gave rise to virion-like particles (VLPs), compared to Pol alone or a Gag-Pol fusion protein that did not form VLPs. The Gag-Pol fusion protein induced the most  
25 broad and potent CTL responses to Gag and Pol in DNA-vaccinated mice, and this immunogen also readily elicited an antibody response to HIV-1 Gag and Pol determinants. Through its ability to induce broad CTL and antibody responses, this Gag-Pol fusion protein or a related derivative is envisioned to be a useful component of an AIDS vaccine.

### Brief Description of the Drawings

**Figure A. Comparison of GP2 with the structures of viral and cellular membrane fusion proteins.**

5 (A) Recombinant Ebola Zaire GP2, (B) Recombinant Mo-55 from the TM subunit of MoMuLv, (C) Low pH-treated HA2 from influenza virus, (D) Recombinant, proteolysis-resistant core of HIV-1 gp41, (E) Recombinant SIV gp41, NMR structure, (F) Recombinant core coiled segments of the SNARES syntaxin 1-A, synaptobrevin-II, and SNAP-25B. Weissenhorn et al., 1998, Molecular Cell, 2,  
10 605-616.

**Figure I. Schematic representation of functional domains and mutations in HIV-1 Env glycoproteins.**

15 Full-length envelope polypeptide, gp160, with the indicated features based on the amino acid residues of HXB2 is shown (top). Functional domains include the gp120/gp41 cleavage site (residues 510/511), the fusion domain (512-527), the two heptad repeats (546-579 and 628-655), the transmembrane domain (684-705), and the cytoplasmic domain (706-856). The mutant forms of the envelope proteins are shown below the structure of gp160. COOH deletions were introduced that  
20 terminate the envelope protein at positions 752, 704, or 680 to produce gp150, gp145, or gp140, respectively. Two internal deletions that removed the cleavage site, the fusion domain, and the region between the two heptad repeats were introduced into gp160, gp150, gp145, and gp140. A further deletion in the COOH-terminal region at position 592 removed the second heptad repeat, the  
25 transmembrane domain, and the interspace region to produce gp128ΔCFI. To disrupt potential glycosylation sites, asparagine (N) residues at eleven positions (88, 156, 160, 197, 230, 234, 241, 262, 276, 289, and 295) were replaced with aspartic acid (D) residues in both gp160 and gp150. Versions of both gp160 and gp150 were created with a total of 17 mutated glycosylation sites by including six additional N  
30 to D substitutions at positions 332, 339, 356, 386, 392, and 448.



**Figure II. Comparison of the expression of the HIV-1 gp160 with codon-optimized gp160.**

A. Expression of plasmids encoding Rev-dependent and Rev-independent codon-modified gp160. Upper panel: expression of Rev-dependent viral gp160 (left) and codon-modified gp160 (right) in transfected 293 cells. Lower panel: comparable expression of  $\beta$ -actin in these transfected cells.

B-D. Expression of mutant CXCR4- and CCR5-tropic HIV Env glycoproteins with COOH-terminal truncations (panels A and B, respectively). CXCR4-tropic envelope proteins containing mutant glycosylation sites and mutant functional domains are shown (panels C and D, respectively). The indicated proteins were detected by immunoblotting as above. Cell lysates produced by transfection with vector containing no insert were used as controls (first lane in each panel).

**Figure III. Cytotoxicity of full-length gp160 is eliminated by deletion of the COOH-terminal cytoplasmic domain.**

Cell rounding and detachment was not observed in control-transfected 293 cells (A), in contrast to full-length gp160 (B) and to a lesser extent in cells transfected with gp150 (C), in contrast to gp145 (D) or gp140 (E).

**Figure IV. Expression of soluble gp140 $\Delta$ CFI HIV-1 envelope variant.**

Immunoprecipitation and Western blot analysis of supernatants from the indicated transfected cells.

**Figure V. Antibody response against HIV-1 envelope proteins in DNA immunized mice.**

A. Comparison of the antibody response in mice immunized with gp140 ( $\Delta$ CFI) or other Env plasmid expression vectors. Sera were collected 2 weeks after the last immunization and used to immunoprecipitate codon-altered gp160 from lysates of transfected 293 cells. The quantitation of the immunoprecipitated gp160 was done as described in Figure 5B. The average of the normalized data has been presented as a bar diagram.

B. Antibody responses in mice immunized with different mutant Env expression vectors. Antisera from immunized mice were diluted in IP buffer and 1



5  $\mu$ l of each diluted serum was used to immunoprecipitate codon-altered HIV-1 gp160 from lysates of transfected 293 cells as described in Figure 3A. The gels were scanned and the intensity of the gp160 band was determined by densitometry using the program Image Quant and presented relative to the intensity of gp160 immunoprecipitated with positive control sera (rabbit anti-gp160), which was used to normalize data between experiments. These data are presented graphically to facilitate comparison among groups.

10 C. Antibody responses in mice immunized with gp140 or gp140 ( $\Delta$ CFI) were determined by immunoprecipitation and Western blotting. Animals received two booster doses (100  $\mu$ g) of the same plasmid, two weeks apart. Sera (1  $\mu$ l) collected 2 weeks after the last immunization was used to immunoprecipitate codon-optimized HIV-1 gp160 from lysates of transfected 293 cells containing 400  $\mu$ g of total protein. Each lane corresponds to the sera from an animal immunized with either the control vector (lanes 1 and 2), CXCR4-tropic gp140 (lanes 3-6), or  
15 plasmid that expresses gp140 with the indicated mutant functional domains (lanes 7-10). A mouse monoclonal antibody to gp160 (HIV-1 V3 Monoclonal (ITB-V3-13), NIH AIDS Reagent Program) was used as a positive control (lane 11).

20 **Figure VI. CTL response against HIV-1 envelope proteins in DNA immunized mice.**

The CTL response to CXCR4-tropic Env and indicated deletion mutants is shown (A). The CTL responses to CXCR4-tropic envelope with glycosylation site and  $\Delta$ CFI mutations are shown (B and C, respectively). Dependence of CTL activity on CD8 cells was shown by magnetic bead depletion using the indicated  
25 representative immunogens (D). Spleen cells were isolated from immunized mice two weeks after the final immunization and stimulated *in vitro* with irradiated cells expressing gp160 with addition of hIL2 (5 U/ml) at day 4. The cytolytic activity of the restimulated spleen cells was tested after 7 days against V3 peptide-pulsed BC10ME cells. Similar findings were observed with target cells that stably express  
30 full-length Env.

**Figure VII. Schematic representation of HIV-1 Gag-Pol expression constructs.**

The protein sequences of Gag (amino acids 1-432) from HXB2 (GenBank accession number K03455) and Pol (amino acids 3-1003) from NL4-3 (GenBank

accession number M19921) were used to create a synthetic version of hGag-Pol using codons found in human cells. hGag-Pol $\Delta$ FS $\Delta$ Pr was made by modification of the frame shift site (FS) and inactivation of protease. For hPol, 432 amino acids were deleted from the NH<sub>2</sub>-terminal region of hGag-Pol and addition of an ATG codon. hGag was made by deletion of 925 amino acids from the COOH-terminal region of hGag-Pol. hGag-Pol, hGag-Pol $\Delta$ FS $\Delta$ Pr, hPol and hGag are expressed from the pNGVL-3 vector backbone.

**Figure VIII. HIV-1 Gag-Pol expression in transfected 293T cells and stably transfected CT26 and BC10 ME cells.**

Cell lysates from 293T cells transfected with pCMV  $\Delta$ R8.2 containing viral Gag-Pol (vGag-Pol) (Naldini, L et al., 1996, Science, 272:263-267), pNGVL-hGag, hPol, hGag-Pol $\Delta$ FS $\Delta$ Pr and hGag-Pol were separated by 4-15% gradient SDS-PAGE, transferred to nitrocellulose filters, and analyzed by immunoblotting with (A) human anti HIV-1-IgG, (B) monoclonal anti-p24, and (C) rabbit anti-RT. (D) Cell lysates from CT26 and BC10ME cells stably transduced with either hGag or hPol were analyzed with human anti HIV-1-IgG.

**Figure IX. Transmission electron microscopy of HIV-1 immature virus-like particles (VLP) produced by transfected 293T.**

Cells were transfected with pNGVL-hGag 48 hours prior to harvesting and fixing (magnification 25,000X).

**Figure X. Gag or Pol specific CTL response mediated by CD8 positive cells in immunized mice.**

Two weeks after mice were immunized with a control vector, hGag, hPol, hGag-Pol $\Delta$ FS $\Delta$ Pr, and hGag-Pol, splenic cells were harvested and sensitized with naïve mouse splenic cells pulsed with Gag or Pol peptides. One week later, effector cells were tested for cytolytic activity in a 5-h <sup>51</sup>Cr release assay using <sup>51</sup>Cr-labeled BC10ME target cells that were pulsed for 2 hours with either (A) HIV-1 Gag peptides, or (B) HIV-1 Pol peptides. (C) CD4+ or CD8+ lymphocytes were depleted from splenic cells of immunized mice with anti-mouse- CD4+ or CD8+ Dynal beads according to the manufacturer's instructions.

**Figure XI. Gag or Pol specific CTL response mediated by CD8 positive cells in immunized mice using stable expressing cell lines as target cells.**

Two weeks after immunization in mice, splenic cells were harvested and sensitized with naïve mouse splenic cells pulsed with Gag or Pol peptides. One week later, effector cells were tested for cytolytic activity in a 5-h <sup>51</sup>Cr release assay using <sup>51</sup>Cr-labeled BC10ME target cells expressing either (A) HIV-1 Gag or (B) Pol protein.

**Figure XII. HIV-1 p24 antibody ELISA assays, HIV-1 immunoblotting and immunoprecipitation Western blotting.**

(A) An HIV-1 p24 antibody ELISA assay was performed by coating 96-well plates with 50 µl of purified recombinant HIV-1<sub>IIIB</sub> p24 antigen at a concentration of 2 µg/ml in PBS buffer, pH 7.4.

(B) HIV-1 immunoblotting of strips containing HIV-1 proteins were incubated with pooled mouse sera at a dilution of 1:25. Bands were visualized using the ECL western blotting detection reagent.

(C) Immunoprecipitation and Western blotting of hPol gene-transfected 293T cell lysates three days after transfection with RIPA buffer. The pooled mouse serum was diluted with IP buffer. After adding 10 µg of the cell lysate containing HIV-1 Pol protein, the reactions were incubated overnight on a rotator at 4°C. The next day, 250 µl of Protein G and A Sepharose beads (10%V/V in IP buffer) were added, and the reactions were incubated on a rotator for 2 hours at 4°C. The reactions were washed 4x with IP buffer, re-suspended with 30 µl of 1X sample buffer, and then loaded onto SDS-PAGE. The reactions were transferred to an Immobilon P membrane, and then incubated with anti HIV-1-IgG. Bands were visualized using the ECL Western blotting detection reagent.

Table 1

Env Plasmids				SEQ ID NO	Figure
Plasmid	Plasmid Name/Description	Plasmid Map Name			
2100	pVR1012x/s R5gp139-Nef (delta) MHC (delta) CD4/h	pVR1012x/s R5gp139-Nef delta MHC delta CD4/h		1	1
2200	pVR1012x/s R5gp157-Nef (delta) MHC (delta) CD4/h	pVR1012x/s R5gp157-Nef delta MHC delta CD4/h		2	2
2300	pVR1012x/s X4 gp139-Nef delta MHC delta CD4/h	pVR1012x/s X4gp139-Nef delta MHC delta CD4/h		3	3
2302	pVR1012x/s X4gp130-Nef/h	pVR1012x/s X4gp130-Nef/h		4	4
2400	pVR1012x/s X4gp157-Nef delta MHC delta CD4/h	pVR1012x/s X4gp157-Nef delta MHC delta CD4/h		5	5
2700	pVR1012x/s X4gp140/h	pVR1012x/s X4gp140/h		6	6
2701	pVR1012x/s X4gp140 ΔCFI/h OR pVR1012x/s X4gp140(delta F/CL del H IS)/h	pVR1012x/s X4gp140(delta F/CL del H IS)/h		7	7
2702	pVR1012x/s X4gp128 ΔCFI/h OR pVR1012x/s X4gp128(delta F/CL)/h	pVR1012x/s X4gp128(delta F/CL)/h		8	8
2706	pVR1012x/s X4gp145/h	pVR1012x/s X4gp145/h		9	9
2707	pVR1012x/s X4gp145 ΔCFI/h OR pVR1012x/s X4gp145(delta F/CL del H IS)/h	pVR1012x/s X4gp145(delta F/CL del H IS)/h		10	10
2800	pVR1012x/s R5gp140/h	pVR1012x/s R5gp140/h		11	11
2801	pVR1012x/s R5gp140 ΔCFI/h OR pVR1012x/s R5gp140(delta F/CL del H IS)/h	pVR1012x/s R5gp140(delta F/CL del H IS)/h		12	12
2804	pVR1012x/s R5gp145/h	pVR1012x/s R5gp145/h		13	13
2805	pVR1012x/s R5gp145 ΔCFI/h OR pVR1012x/s R5gp145(delta F/CL del H IS)/h	pVR1012x/s R5gp145(delta F/CL del H IS)/h		14	14
2810	pVR1012x/s R5gp140delC1(delta CFI)/h	pVR1012x/s R5gp140delC1(delta CFI)/h		15	15
2811	pVR1012x/s R5gp140delC2(delta CFI)/h	pVR1012x/s R5gp140delC2(delta CFI)/h		16	16
2812	pVR1012x/s R5gp140delC3(delta CFI)/h	pVR1012x/s R5gp140delC3(delta CFI)/h		17	17
2813	pVR1012x/s R5gp140delC4 (delta CFI)/h	pVR1012x/s R5gp140delC4(delta CFI)/h		18	18
2814	pVR1012x/s R5gp140delC5 (delta CFI)/h	pVR1012x/s R5gp140delC5(delta CFI)/h		19	19
2820	pVR1012x/s R5gp140(dCFI)/dV1	pVR1012x/s R5gp140(dCFI)dV1/h		20	20
2821	pVR1012x/s R5gp140(dCFI)/dV2	pVR1012x/s R5gp140(dCFI)dV2/h		21	21
2822	pVR1012x/s R5gp140(dCFI)/dV3	pVR1012x/s R5gp140(dCFI)dV3/h		22	22
2823	pVR1012x/s R5gp140(dCFI)/dV4	pVR1012x/s R5gp140(dCFI)dV4/h		23	23
2824	pVR1012x/s R5gp140(dCFI)/dV12	pVR1012x/s R5gp140(dCFI)dV12/h		24	24
2825	pVR1012x/s R5gp140(dCFI)/dV13	pVR1012x/s R5gp140(dCFI)dV13/h		25	25
2826	pVR1012x/s R5gp140(dCFI)/dV14	pVR1012x/s R5gp140(dCFI)dV14/h		26	26
2827	pVR1012x/s R5gp140(dCFI)/dV23	pVR1012x/s R5gp140(dCFI)dV23/h		27	27
2828	pVR1012x/s R5gp140(dCFI)/dV24	pVR1012x/s R5gp140(dCFI)dV24/h		28	28
2829	pVR1012x/s R5gp140(dCFI)/dV34	pVR1012x/s R5gp140(dCFI)dV34/h		29	29
2830	pVR1012x/s R5gp140(dCFI)/dV123	pVR1012x/s R5gp140(dCFI)dV123/h		30	30



Env Plasmids		Plasmid Name/Description			Plasmid Map Name		SEQ ID NO	Figure
Plasmid		Plasmid Name/Description			Plasmid Map Name		SEQ ID NO	Figure
2831		pVR1012x/s R5gp140(dCFI)/dv124			pVR1012x/s R5gp140(dCFI)/dv124/h		31	31
2832		pVR1012x/s R5gp140(dCFI)/dv134			pVR1012x/s R5gp140(dCFI)/dv134/h		32	32
2833		pVR1012x/s R5gp140(dCFI)/dv234			pVR1012x/s R5gp140(dCFI)/dv234/h		33	33
2834		pVR1012x/s R5gp140(dCFI)/dv1234			pVR1012x/s R5gp140(dCFI)/dv1234/h		34	34
2835		pAdApt R5gp140(dCFI)/dv1			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv1/h		35	35
2836		pAdApt R5gp140(dCFI)/dv2			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv2/h		36	36
2837		pAdApt R5gp140(dCFI)/dv3			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv3/h		37	37
2838		pAdApt R5gp140(dCFI)/dv4			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv4/h		38	38
2839		pAdApt R5gp140(dCFI)/dv12			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv12/h		39	39
2840		pAdApt R5gp140(dCFI)/dv13			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv13/h		40	40
2841		pAdApt R5gp140(dCFI)/dv14			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv14/h		41	41
2842		pAdApt R5gp140(dCFI)/dv23			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv23/h		42	42
2843		pAdApt R5gp140(dCFI)/dv24			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv24/h		43	43
2844		pAdApt R5gp140(dCFI)/dv34			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv34/h		44	44
2845		pAdApt R5gp140(dCFI)/dv123			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv123/h		45	45
2846		pAdApt R5gp140(dCFI)/dv124			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv124/h		46	46
2847		pAdApt R5gp140(dCFI)/dv134			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv134/h		47	47
2848		pAdApt R5gp140(dCFI)/dv234			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv234/h		48	48
2849		pAdApt R5gp140(dCFI)/dv1234			pAdApt CMV TbGH (+) R5gp140(dCFI)/dv1234/h		49	49
2850		pVR1012x/s R5gp145delC1 (delCFI)/h			pVR1012x/sR5gp145delC1 (delCFI)/h		50	50
2851		pVR1012x/s R5gp145delC2 (delCFI)/h			pVR1012x/sR5gp145delC2 (delCFI)/h		51	51
2852		pVR1012x/s R5gp145delC3 (delCFI)/h			pVR1012x/sR5gp145delC3 (delCFI)/h		52	52
2853		pVR1012x/s R5gp145delC4 (delCFI)/h			pVR1012x/sR5gp145delC4 (delCFI)/h		53	53
2854		pVR1012x/s R5gp145delC5 (delCFI)/h			pVR1012x/sR5gp145delC5 (delCFI)/h		54	54
2860		pVR1012x/s R5gp145(dCFI)/h/dv1			pVR1012x/s R5 gp145(dCFI)/dv1/h		55	55
2861		pVR1012x/s R5gp145(dCFI)/h/dv2			pVR1012x/s R5gp145(dCFI)/dv2/h		56	56
2862		pVR1012x/s R5gp145(dCFI)/h/dv3			pVR1012x/s R5gp145(dCFI)/dv3/h		57	57
2863		pVR1012x/s R5gp145(dCFI)/h/dv4			pVR1012x/s R5gp145(dCFI)/dv4/h		58	58
2864		pVR1012x/s R5gp145(dCFI)/h/dv12			pVR1012x/s R5gp145(dCFI)/dv12/h		59	59
2865		pVR1012x/s R5gp145(dCFI)/h/dv13			pVR1012x/s R5gp145(dCFI)/dv13/h		60	60
2866		pVR1012x/s R5gp145(dCFI)/h/dv14			pVR1012x/s R5gp145(dCFI)/dv14/h		61	61
2867		pVR1012x/s R5gp145(dCFI)/h/dv23			pVR1012x/s R5gp145(dCFI)/dv23/h		62	62

Plasmid		Env Plasmids		Plasmid Name/Description		Plasmid Map Name		SEQ ID NO	Figure
2868				pVR1012x/s R5gp145(dCFI)/h/dV24		pVR1012x/s R5gp145(dCFI)dv24/h		63	63
2869				pVR1012x/s R5gp145(dCFI)/h/dV34		pVR1012x/s R5gp145(dCFI)dv34/h		64	64
2870				pVR1012x/s R5gp145(dCFI)/h/dV134		pVR1012x/s R5gp145(dCFI)dv134/h		65	65
2871				pVR1012x/s R5gp145(dCFI)/h/dV234		pVR1012x/s R5gp145(dCFI)dv234/h		66	66
2872				pVR1012x/s R5gp145(dCFI)/h/dV123		pVR1012x/s R5gp145(dCFI)dv123/h		67	67
2873				pVR1012x/s R5gp145(dCFI)/h/dV124		pVR1012x/s R5gp145(dCFI)dv124/h		68	68
2874				pVR1012x/s R5gp145(dCFI)/h/dV1234		pVR1012x/s R5gp145(dCFI)dv1234/h		69	69
2900				pVR1012x/s R5gp150/h		pVR1012x/s R5gp150/h		70	70
3000				pVR1012x/s R5gp160/h		pVR1012x/s R5gp160/h		71	71
3200				pVR1012x/s X4gp150/h		pVR1012x/s X4gp150/h		72	72
3201				pVR1012x/s X4gp150 ΔCFI/h OR pVR1012x/s X4gp150(dcl F/CL del H IS)/h		pVR1012x/s X4gp150(dcl F/CL del H IS)/h		73	73
3202				pVR1012x/s X4gp150 Δgly/h		pVR1012x/s X4gp150 delta gly		74	74
3203				pVR1012x/s X4gp150 AB Δgly/h		pVR1012x/s X4gp150 AB (delta)gly/h		75	75
3300				pVR1012x/s X4gp160/h		pVR1012x/s X4gp160/h		76	76
3301				pVR1012x/s X4gp160 ΔCFI/h OR pVR1012x/s X4gp160(dcl F/CL del H IS)/h		pVR1012x/s X4gp160(dcl F/CL del H IS)/h		77	77
3400				pVR1012x/s X4gp160 Δgly/h		pVR1012x/s X4gp160 delta gly		78	78
3401				pVR1012x/s X4gp160 AB Δgly/h OR pVR1012x/s X4gp160AB mut Δgly/h		pVR1012x/s X4gp160 AB Dgly/h		79	79
3500				pVR1012x/s Nef/h		pVR1012x/s Nef/h		80	80
3600				pVR1012x/s NefDMHCDCD4/h		pVR1012x/s Nef delta MHC delta CD4/h		81	81
3700				pVR1012x/s NefDCD4/h		pVR1012x/s Nef delta CD4/h		82	82
3800				pVR1012x/s NefDMHC/h		pVR1012x/s Nef delta MHC/h		83	83
5200				pVR1012 x/s 89.6Pgp128(dcl F/CL)/h		pVR1012 x/s 89.6Pgp128(dcl F/CL)/h		84	84
5201				R5 Clade 89.6P gp140 ΔCFI/h		pVR1012x/s 89.6Pgp140(dcl F/CL del H IS)/h		85	85
5202				pVR1012 x/s 89.6Pgp145(dcl F/CL del H IS)/h		pVR1012 x/s 89.6Pgp145(dcl F/CL del H IS)/h		86	86
5203				pVR1012 x/s 89.6Pgp160/h		pVR1012 x/s 89.6Pgp160/h		87	87
5300				R5 Clade C gp140 ΔCFI/h OR pVR1012x/s R5 (cladeC)gp140(dcl F/CL del H IS)/h		pVR1012x/s CladeC(R5)gp140(dcl F/CL del H IS)/h		88	88
5301				pVR1012x/s R5(cladeC)gp145(dcl F/CL del H IS)/h		pVR1012x/s CladeC(R5)gp145(dcl F/CL del H IS)/h		89	89
5303				pVR1012x/s R5gp145 CladeC (Brazil) delCFI/h		pVR1012x/s R5gp145 CladeC (Brazil) delCFI/h		90	90
5304				pVR1012x/s R5(clade A)gp140(dcl F/CL del H IS)/h		pVR1012x/s R5gp140CladeA (dCFI)/h		91	91
5305				pVR1012x/s R5(clade A)gp145(dcl F/CL del H IS)/h		pVR1012x/s R5gp145CladeA (dCFI)/h		92	92
5306				pVR1012x/s R5(clade E)gp140(dcl F/CL del H IS)/h		pVR1012x/s R5gp140 Clade E (dCFI)/h		93	93
5307				pVR1012x/s R5(clade E)gp145(dcl F/CL del H IS)/h		pVR1012x/s R5 gp145 Clade E (dCFI)/h		94	94



Env Plasmids				SEQ ID NO	Figure
Plasmid	Plasmid Name/Description	Plasmid Map Name			
5308	pVR1012x/s R5(clade C South African)gp140(del F/CL del H IS)/h	pVR1012x/s R5gp140 Clade C(SA) (dCFI)/h		95	95
5309	pVR1012x/s R5(clade C South African)gp145(del F/CL del H IS)/h	pVR1012x/s R5gp145 CladeC(SA) (dCFI)/h		96	96
5350	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1	pVR1012x/s R5gp140CladeC (Brazil) dCFI/dv1/h		97	97
5351	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV12	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv12/h		98	98
5352	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV123	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv123/h		99	99
5353	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1234	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv1234/h		100	100
5354	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV124	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv124/h		101	101
5355	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV13	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv13/h		102	102
5356	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV134	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv134/h		103	103
5357	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV14	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv14/h		104	104
5358	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV2	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv2/h		105	105
5359	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV23	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv23/h		106	106
5360	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV234	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv234/h		107	107
5361	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV24	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv24/h		108	108
5362	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV3	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv3/h		109	109
5363	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV34	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv34/h		110	110
5364	pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV4	pVR1012x/s R5gp140CladeC (Brazil) (dCFI)dv4/h		111	111
5365	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv1/h		112	112
5366	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV12	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv12/h		113	113
5367	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV123	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv123/h		114	114
5368	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1234	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv1234/h		115	115
5369	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV124	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv124/h		116	116
5370	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV13	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv13/h		117	117
5371	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV134	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv134/h		118	118
5372	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV14	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv14/h		119	119
5373	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV2	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv2/h		120	120
5374	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV23	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv23/h		121	121
5375	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV234	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv234/h		122	122
5376	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV24	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv24/h		123	123
5377	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV3	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv3/h		124	124
5378	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV34	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv34/h		125	125
5379	pVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV4	pVR1012x/s R5gp145CladeC (Brazil) (dCFI)dv4/h		126	126

Env Plasmids			
Plasmid	Plasmid Name/Description	Plasmid Map Name	SEQ ID NO
5500	pVR1012x/s R5(SA-C)gp140(dCFI) dV1/h	pVR1012x/s R5 gp140(dCFI) SA dv1/h	127
5501	pVR1012x/s R5(SA-C)gp140(dCFI) dV12/h	pVR1012x/s R5 gp140(dCFI) SA dv12/h	128
5502	pVR1012x/s R5(SA-C)gp140(dCFI) dV123/h	pVR1012x/s R5 gp140(dCFI) SA dv123/h	129
5503	pVR1012x/s R5(SA-C)gp140(dCFI) dV1234/h	pVR1012x/s R5 gp140(dCFI) SA dv1234/h	130
5504	pVR1012x/s R5(SA-C)gp140(dCFI) dV124/h	pVR1012x/s R5 gp140(dCFI) SA dv124/h	131
5505	pVR1012x/s R5(SA-C)gp140(dCFI) dV13/h	pVR1012x/s R5 gp140(dCFI) SA dv13/h	132
5506	pVR1012x/s R5(SA-C)gp140(dCFI) dV134/h	pVR1012x/s R5 gp140(dCFI) SA dv134/h	133
5507	pVR1012x/s R5(SA-C)gp140(dCFI) dV14/h	pVR1012x/s R5 gp140(dCFI) SA dv14/h	134
5508	pVR1012x/s R5(SA-C)gp140(dCFI) dV2/h	pVR1012x/s R5 gp140(dCFI) SA dv2/h	135
5509	pVR1012x/s R5(SA-C)gp140(dCFI) dV23/h	pVR1012x/s R5 gp140(dCFI) SA dv23/h	136
5510	pVR1012x/s R5(SA-C)gp140(dCFI) dV234/h	pVR1012x/s R5 gp140(dCFI) SA dv234/h	137
5511	pVR1012x/s R5(SA-C)gp140(dCFI) dV24/h	pVR1012x/s R5 gp140(dCFI) SA dv24/h	138
5512	pVR1012x/s R5(SA-C)gp140(dCFI) dV3/h	pVR1012x/s R5 gp140(dCFI) SA dv3/h	139
5513	pVR1012x/s R5(SA-C)gp140(dCFI) dV34/h	pVR1012x/s R5 gp140(dCFI) SA dv34/h	140
5514	pVR1012x/s R5(SA-C)gp140(dCFI) dV4/h	pVR1012x/s R5 gp140(dCFI) SA dv4/h	141
5515	pVR1012x/s R5(SA-C)gp145(dCFI) dV1/h	pVR1012x/s R5 gp145(dCFI) SA dv1/h	142
5516	pVR1012x/s R5(SA-C)gp145(dCFI) dV12/h	pVR1012x/s R5 gp145(dCFI) SA dv12/h	143
5517	pVR1012x/s R5(SA-C)gp145(dCFI) dV123/h	pVR1012x/s R5 gp145(dCFI) SA dv123/h	144
5518	pVR1012x/s R5(SA-C)gp145(dCFI) dV1234/h	pVR1012x/s R5 gp145(dCFI) SA dv1234/h	145
5519	pVR1012x/s R5(SA-C)gp145(dCFI) dV2/h	pVR1012x/s R5 gp145(dCFI) SA dv2/h	146
5520	pVR1012x/s R5(SA-C)gp145(dCFI) dV23/h	pVR1012x/s R5 gp145(dCFI) SA dv23/h	147
5521	pVR1012x/s R5(SA-C)gp145(dCFI) dV234/h	pVR1012x/s R5 gp145(dCFI) SA dv234/h	148
5522	pVR1012x/s R5(SA-C)gp145(dCFI) dV24/h	pVR1012x/s R5 gp145(dCFI) SA dv24/h	149
5523	pVR1012x/s R5(SA-C)gp145(dCFI) dV3/h	pVR1012x/s R5 gp145(dCFI) SA dv3/h	150
5524	pVR1012x/s R5(SA-C)gp145(dCFI) dV34/h	pVR1012x/s R5 gp145(dCFI) SA dv34/h	151
5525	pVR1012x/s R5(SA-C)gp145(dCFI) dV4/h	pVR1012x/s R5 gp145(dCFI) SA dv4/h	152
5526	pVR1012x/s R5(SA-C)gp145(dCFI) dV13/h	pVR1012x/s R5 gp145(dCFI) SA dv13/h	153
5527	pVR1012x/s R5(SA-C)gp145(dCFI) dV134/h	pVR1012x/s R5 gp145(dCFI) SA dv134/h	154
5528	pVR1012x/s R5(SA-C)gp145(dCFI) dV124/h	pVR1012x/s R5 gp145(dCFI) SA dv124/h	155
5529	pVR1012x/s R5(SA-C)gp145(dCFI) dV14/h	pVR1012x/s R5 gp145(dCFI) SA dv14/h	156

Gag/Pol Plasmids				
Plasmid	Plasmid Name/Description	Plasmid Map Name	SEQ ID NO	Figure
3900	pVR1012x/s HIV Gag /h	pVR1012x/s Gag/h	157	157
3901	pVR1012x/s SIV Gag /h	pVR1012x/s SIV Gag/h	158	158
4000	pVR1012x/s HIV Gag-Pol ΔFS ΔPR /h OR pVR1012x/s Gag-Pol/h	pVR1012x/s Gag-Pol deltaFS deltaPR/h	159	159
4001	pVR1012x/s SIV Gag-Pol /h	pVR1012x/s SIV Gag-Pol/h	160	160
4100	pVR1012x/s HIV Pol /h	pVR1012x/s Pol/h	161	161
4101	pVR1012x/s SIV Pol /h	pVR1012x/s SIV Pol/h	162	162
4200	pVR1012x/s HIV Gag-Pol /h	pVR1012x/s Gag-Pol/h	163	163
4300	pVR1012x/s HIV Gag-Pol ΔIN /h OR pVR1012 Gag-Pol(d delta RT delta IN)/h	pVR1012x/s Gag-Pol(fs)RT(-)IN(-)	164	164
4301	pVR1012x/s-Gag(FS)-Pol-delta RT IN-IRES-R5 gp157-Nef	pVR1012x/s-Gag(FS)-Pol-delta RT IN-IRES-R5 gp157-Nef	165	165
4302	pVR1012x/s HIV gag-Pol ΔFS ΔPR ΔRT ΔIN /h	pVR1012x/s Gag(deltaFS)Pol (delta PR delta RT delta IN)/h	166	166
4303	pVR1012x/s SIV Gag-Pol ΔFS /h OR pVR1012 SIV Gag(deltaFS)Pol(delta PR delta RT delta IN)/h	pVR1012x/s SIV Gag(deltaFS)-Pol/h	167	167
4304	pVR1012 Gag (deltaFS)Pol delta PR delta RT delta IN/h	pVR1012 Gag-C(deltaFS) Pol (deltaPR deltaRT deltaIN)/h	168	168
4305	pVR1012 Gag-A(deltaFS)Pol(delta PR delta RT delta IN)/h	pVR1012 Gag-A(deltaFS) Pol (deltaPR delta RT deltaIN)/h	169	169
4306	pVR1012 Gag(deltaFS)Pol delta PR delta RT delta IN/Nef/h	pVR1012 Gag(deltaFS)Pol deltaPR deltaRT deltaIN/Nef/h	170	170
4308	pVR1012 Gag (deltaFS) Pol deltaPR deltaRT deltaIN deltaMyr/h	pVR1012 Gag(deltaFS)Pol deltaPR deltaRT deltaIN delta Myr /h	171	171
4309	pVR1012 Gag (deltaFS) Pol deltaPR deltaRT deltaIN delta Myr/Nef/h	pVR1012x/s Gag(deltaFS)Pol deltaPR deltaRT deltaIN delta Myr/Nef/h	172	172
4310	pVR1012 Nef Gag (delta fs) (delta Myr) Pol (delta PR delta RT delta IN)/h	pVR1012 Nef Gag(deltaFS)(delta Myr) Pol (delta PR delta RT delta IN)/h	173	173
4311	pVR1012 Gag-C(deltaFS)Pol(delta PR delta RT delta IN) Nef/h	pVR1012 Gag-C(delta FS) Pol(delta PR delta RT delta IN) Nef/h	174	174
4312	Gag (delta fs) (delta Myr) Nef Pol ΔPRΔRTΔIN/h	pVR1012 Gag (delta FS) (delta Myr) Nef Pol (delta PR delta RT delta IN)/h	175	175
4313	pVR1012 Gag Clade A (delta fs)Pol(Δ PR Δ RT Δ IN)/h	pVR1012 Gag-A (delta FS) Pol (deltaPR deltaRT deltaIN) Nef/h	176	176



### Detailed Description of the Invention

The present invention provides modifications of HIV Env, Gag, Pol, and Nef that enhance immunogenicity for genetic immunization. Both HIV and SIV are genetically related members of the lentivirus genus of the *Retroviridae* family. Lentivirus isolates from humans are grouped into one of two types, designated HIV-1 and HIV-2. A classification scheme recognizes nine subtypes (clades) of HIV-1 (A through I) and five subtypes of HIV-2 (A through E). A compendium of HIV and SIV sequence information is found in a database prepared by Myers et al., Los Alamos, NM: Los Alamos National Laboratory.

#### 10 Nucleic Acid Molecules

As indicated herein, nucleic acid molecules of the present invention may be in the form of RNA or in the form of DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

25 Nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) of a wild-type HIV gene; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode an ORF of a wild-type HIV polypeptide. Of course, the genetic code is well known in the art. Degenerate variants optimized for human codon usage are preferred.

30 In another aspect, the invention provides a nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above. By "stringent hybridization conditions" is intended overnight incubation at 42 degree C in a solution comprising: 50% formamide, 5 times SSC (750 mM NaCl, 75

mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 times Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 times SSC at about 65 degree C.

5 By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 nt of the reference polynucleotide.

10 By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide. Of course, a polynucleotide which hybridizes only to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule  
15 containing a poly T (or U) stretch or the complement thereof (e.g., practically any double-stranded DNA clone).

As indicated herein, nucleic acid molecules of the present invention which encode an HIV polypeptide may include, but are not limited to those encoding the amino acid sequence of the full-length polypeptide, by itself, the coding sequence  
20 for the full-length polypeptide and additional sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro-protein sequence, the coding sequence of the full-length polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences,  
25 such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example, ribosome binding and stability of mRNA; and additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities.

30 The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the HIV protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a genome of an organism. Genes II, Lewin, B., ed., John Wiley & Sons, New

York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions, which may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the HIV polypeptide or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 95% identical, and more preferably at least 96%, 97%, 98% or 99% identical to a nucleotide sequence encoding a polypeptide having the amino acid sequence of a wild-type HIV polypeptide or a nucleotide sequence complementary thereto.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a HIV polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the HIV polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 95%, 96%, 97%, 98% or 99% identical to the reference nucleotide sequence can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). Bestfit uses the local homology algorithm of Smith and Waterman,



Advances in Applied Mathematics 2: 482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

The present application is directed to nucleic acid molecules at least 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequences shown herein in the Sequence Listing which encode a polypeptide having HIV polypeptide activity. By "a polypeptide having HIV activity" is intended polypeptides exhibiting HIV activity in a particular biological assay. For example, Env, Gag, and Pol protein activity can be measured for changes in immunological character by an appropriate immunological assay.

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 95%, 96%, 97%, 98%, or 99% identical to a nucleic acid sequence shown herein in the Sequence Listing will encode a polypeptide "having HIV polypeptide activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having HIV polypeptide activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid).

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990), wherein the authors indicate that proteins are surprisingly tolerant of amino acid substitutions.

### Polypeptides and Fragments

The invention further provides a HIV polypeptide having the amino acid sequence encoded by an open reading frame (ORF) of a wild-type HIV gene, or a peptide or polypeptide comprising a portion thereof (e.g., gp120).

5 It will be recognized in the art that some amino acid sequences of the HIV polypeptides can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

10 Thus, the invention further includes variations of the HIV polypeptide which show substantial HIV polypeptide activity or which include regions of HIV protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions. As indicated, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J. U., et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990).

15 Thus, the fragment, derivative or analog of the polypeptide of the invention may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one  
20 encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which additional amino acids are fused to the mature polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are  
25 deemed to be within the scope of those skilled in the art from the teachings herein.

As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table A).

**Table A**  
**Conservative Amino Acid Substitutions**

<b>Aromatic</b>	Phenylalanine Tryptophan Tyrosine
<b>Ionizable: Acidic</b>	Aspartic Acid Glutamic Acid
<b>Ionizable: Basic</b>	Arginine Histidine Lysine
<b>Nonionizable Polar</b>	Asparagine Glutamine Selenocystine Serine Threonine
<b>Nonpolar (Hydrophobic)</b>	Alanine Glycine Isoleucine Leucine Proline Valine
<b>Sulfur Containing</b>	Cysteine Methionine

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of amino acid substitutions for any given HIV polypeptide will not be more than 50, 40, 30, 20, 10, 5 or 3.

Amino acids in the HIV polypeptides of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as changes in immunological character.

The polypeptides of the present invention are conveniently provided in an isolated form. By "isolated polypeptide" is intended a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention.

5 Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially, from a recombinant host cell or a native source. For example, a recombinantly produced version of the HIV polypeptide can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).

10 The polypeptides of the present invention include a polypeptide comprising a polypeptide shown herein in the Sequence Listing; as well as polypeptides which are at least 95% identical, and more preferably at least 96%, 97%, 98% or 99% identical to those described above and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids.

15 By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of an HIV polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the HIV polypeptide. In other words, to obtain a polypeptide having an amino acid sequence  
20 at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These  
25 alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 95%,  
30 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown herein in the Sequence Listing can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). When using Bestfit or any other  
35 sequence alignment program to determine whether a particular sequence is, for

instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are  
5 allowed.

The polypeptides of the invention may be produced by any conventional means. Houghten, R. A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. Proc. Natl. Acad. Sci. USA 82:5131-5135. This  
10 "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Pat. No. 4,631,211 to Houghten et al. (1986).

The present invention also relates to vectors which include the nucleic acid molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of HIV polypeptides or fragments  
15 thereof by recombinant techniques.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line  
20 and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression  
25 constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate  
30 hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such  
35



as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

5 Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other  
10 suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et  
15 al., *Basic Methods In Molecular Biology* (1986).

The HIV polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin  
20 chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

### 30 **Pharmaceutical Formulations, Dosages, and Modes of Administration**

The compounds of the invention may be administered using techniques well known to those in the art. Preferably, compounds are formulated and administered by genetic immunization. Techniques for formulation and administration may be found in "Remington's Pharmaceutical Sciences", 18<sup>th</sup> ed., 1990, Mack Publishing  
35 Co., Easton, PA. Suitable routes may include parenteral delivery, such as



intramuscular, intradermal, subcutaneous, intramedullary injections, as well as, intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections, just to name a few. For injection, the compounds of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer.

In instances wherein intracellular administration of the compounds of the invention is preferred, techniques well known to those of ordinary skill in the art may be utilized. For example, such compounds may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are effectively delivered into the cell cytoplasm.

Nucleotide sequences of the invention which are to be intracellularly administered may be expressed in cells of interest, using techniques well known to those of skill in the art. For example, expression vectors derived from viruses such as retroviruses, adenoviruses, adeno-associated viruses, herpes viruses, vaccinia viruses, polio viruses, or sindbis or other RNA viruses, or from plasmids may be used for delivery and expression of such nucleotide sequences into the targeted cell population. Methods for the construction of such expression vectors are well known. See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY, and Ausubel et al., 1989, Current Protocols In Molecular Biology, Greene Publishing Associates and Wiley Interscience, NY.

The invention extends to the use of a plasmid for primary immunization (priming) of a host and the subsequent use of a recombinant virus, such as a retrovirus, adenovirus, adeno-associated virus, herpes virus, vaccinia virus, polio virus, or sindbis or other RNA virus, for boosting said host, and vice versa. For example, the host may be immunized (primed) with a plasmid by DNA immunization and receive a boost with the corresponding viral construct, and vice versa. Alternatively, the host may be immunized (primed) with a plasmid by DNA immunization and receive a boost with not the corresponding viral construct but a different viral construct, and vice versa.

With respect to HIV Env, Gag, and Pol, protein sequences of the invention may be used as therapeutics or prophylactics (as subunit vaccines) in the treatment of AIDS or HIV infection. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that includes the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (e.g., the concentration of the test compound which achieves a half-maximal inhibition of viral infection relative to the amount of the event in the absence of the test compound) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography (HPLC).

The compounds of the invention may, further, serve the role of a prophylactic vaccine, wherein the host produces antibodies and/or CTL responses against HIV Env, Gag and Pol, which responses then preferably serve to neutralize HIV viruses by, for example, inhibiting further HIV infection. Administration of the compounds of the invention as a prophylactic vaccine, therefore, would comprise administering to a host a concentration of compounds effective in raising an immune response which is sufficient to elicit antibody and/or CTL responses to HIV Env, Gag, and Pol, and/or neutralize HIV, by, for example, inhibiting HIV ability to infect cells. The exact concentration will depend upon the specific compound to be administered, but may be determined by using standard techniques for assaying the development of an immune response which are well known to those of ordinary skill in the art.

The compounds may be formulated with a suitable adjuvant in order to enhance the immunological response. Such adjuvants may include, but are not limited to mineral gels such as aluminum hydroxide; surface active substances such as lysolecithin, pluronic polyols, polyanions; other peptides; oil emulsions; and  
5 potentially useful human adjuvants such as BCG and *Corynebacterium parvum*.

Adjuvants suitable for co-administration in accordance with the present invention should be ones that are potentially safe, well tolerated and effective in people including QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjumer, PG-026, GSK-1, GcMAF, B-  
10 alethine, MPC-026, Adjuvax, CpG ODN, Betafectin, Alum, and MF59 (see Kim et al., 2000, Vaccine, 18: 597 and references therein).

Other contemplated adjuvants that may be administered include lectins, growth factors, cytokines and lymphokines such as alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), gCSF, gMCSF, TNF, epidermal  
15 growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10 and IL-12.

For all such treatments described above, the exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p. 1).

20 It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administered dose in the management of the viral  
25 infection of interest will vary with the severity of the condition to be treated and the route of administration. The dose and perhaps prime-boost regimen, will also vary according to the age, weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

The pharmacologically active compounds of this invention can be processed  
30 in accordance with conventional methods of galenic pharmacy to produce medicinal agents for administration to patients, e.g., mammals including humans.

The compounds of this invention can be employed in admixture with conventional excipients, i.e., pharmaceutically acceptable organic or inorganic carrier substances suitable for parenteral, enteral (e.g., oral) or topical application  
35 which do not deleteriously react with the active compounds. Suitable

pharmaceutically acceptable carriers include but are not limited to water, salt solutions, alcohols, gum arabic, vegetable oils, benzyl alcohols, polyethylene glycols, gelatine, carbohydrates such as lactose, amylose or starch, magnesium stearate, talc, silicic acid, viscous paraffin, perfume oil, fatty acid monoglycerides and diglycerides, pentaerythritol fatty acid esters, hydroxy methylcellulose, polyvinyl pyrrolidone, etc. The pharmaceutical preparations can be sterilized and if desired mixed with auxiliary agents, e.g., lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, flavoring and/or aromatic substances and the like which do not deleteriously react with the active compounds. They can also be combined where desired with other active agents, e.g., vitamins.

For parenteral application, which includes intramuscular, intradermal, subcutaneous, intranasal, intracapsular, intraspinal, intrasternal, and intravenous injection, particularly suitable are injectable, sterile solutions, preferably oily or aqueous solutions, as well as suspensions, emulsions, or implants, including suppositories. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

For enteral application, particularly suitable are tablets, dragees, liquids, drops, suppositories, or capsules. The pharmaceutical compositions may be prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated



edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate. A syrup, elixir, or the like can be used wherein a sweetened vehicle is employed.

Sustained or directed release compositions can be formulated, e.g., liposomes or those wherein the active compound is protected with differentially degradable coatings, e.g., by microencapsulation, multiple coatings, etc. It is also possible to freeze dry the new compounds and use the lyophilizates obtained, for example, for the preparation of products for injection.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

For topical application, there are employed as non-sprayable forms, viscous to semi-solid or solid forms comprising a carrier compatible with topical application and having a dynamic viscosity preferably greater than water. Suitable formulations include but are not limited to solutions, suspensions, emulsions, creams, ointments, powders, liniments, salves, aerosols, etc., which are, if desired, sterilized or mixed with auxiliary agents, e.g., preservatives, stabilizers, wetting agents, buffers or salts for influencing osmotic pressure, etc. For topical application, also suitable are sprayable aerosol preparations wherein the active ingredient, preferably in combination with a solid or liquid inert carrier material, is packaged in a squeeze bottle or in admixture with a pressurized volatile, normally gaseous propellant, e.g., a freon.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

### Genetic Immunization

Genetic immunization according to the present invention elicits an effective immune response without the use of infective agents or infective vectors. Vaccination techniques which usually do produce a CTL response do so through the use of an infective agent. A complete, broad based immune response is not generally exhibited in individuals immunized with killed, inactivated or subunit vaccines. The present invention achieves the full complement of immune responses in a safe manner without the risks and problems associated with vaccinations that use infectious agents.

According to the present invention, DNA or RNA that encodes a target protein is introduced into the cells of an individual where it is expressed, thus producing the target protein. The DNA or RNA is linked to regulatory elements necessary for expression in the cells of the individual. Regulatory elements for DNA include a promoter and a polyadenylation signal. In addition, other elements, such as a Kozak region, may also be included in the genetic construct.

The genetic constructs of genetic vaccines comprise a nucleotide sequence that encodes a target protein operably linked to regulatory elements needed for gene expression. Accordingly, incorporation of the DNA or RNA molecule into a living cell results in the expression of the DNA or RNA encoding the target protein and thus, production of the target protein.

When taken up by a cell, the genetic construct which includes the nucleotide sequence encoding the target protein operably linked to the regulatory elements may remain present in the cell as a functioning extrachromosomal molecule or it may integrate into the cell's chromosomal DNA. DNA may be introduced into cells where it remains as separate genetic material in the form of a plasmid. Alternatively, linear DNA which can integrate into the chromosome may be introduced into the cell. When introducing DNA into the cell, reagents which promote DNA integration into chromosomes may be added. DNA sequences which are useful to promote integration may also be included in the DNA molecule. Since integration into the chromosomal DNA necessarily requires manipulation of the chromosome, it is preferred to maintain the DNA construct as a replicating or non-replicating extrachromosomal molecule. This reduces the risk of damaging the cell by splicing into the chromosome without affecting the effectiveness of the vaccine. Alternatively, RNA may be administered to the cell. It is also contemplated to

provide the genetic construct as a linear minichromosome including a centromere, telomeres and an origin of replication.

5 The necessary elements of a genetic construct of a genetic vaccine include a nucleotide sequence that encodes a target protein and the regulatory elements necessary for expression of that sequence in the cells of the vaccinated individual. The regulatory elements are operably linked to the DNA sequence that encodes the target protein to enable expression.

10 The molecule that encodes a target protein is a protein-encoding molecule which is translated into protein. Such molecules include DNA or RNA which comprise a nucleotide sequence that encodes the target protein. These molecules may be cDNA, genomic DNA, synthesized DNA or a hybrid thereof or an RNA molecule such as mRNA. Accordingly, as used herein, the terms "DNA construct", "genetic construct" and "nucleotide sequence" are meant to refer to both DNA and RNA molecules.

15 The regulatory elements necessary for gene expression of a DNA molecule include: a promoter, an initiation codon, a stop codon, and a polyadenylation signal. In addition, enhancers are often required for gene expression. It is necessary that these elements be operable in the vaccinated individual. Moreover, it is necessary that these elements be operably linked to the nucleotide sequence that encodes the target protein such that the nucleotide sequence can be expressed in the cells of a vaccinated individual and thus the target protein can be produced.

Initiation codons and stop codons are generally considered to be part of a nucleotide sequence that encodes the target protein. However, it is necessary that these elements are functional in the vaccinated individual.

25 Similarly, promoters and polyadenylation signals used must be functional within the cells of the vaccinated individual.

30 Examples of promoters useful to practice the present invention, especially in the production of a genetic vaccine for humans, include but are not limited to promoters from Simian Virus 40 (SV40), Mouse Mammary Tumor Virus (MMTV) promoter, Human Immunodeficiency Virus (HIV) such as the HIV Long Terminal Repeat (LTR) promoter, Moloney virus, ALV, Cytomegalovirus (CMV) such as the CMV immediate early promoter, Epstein Barr Virus (EBV), Rous Sarcoma Virus (RSV) as well as promoters from human genes such as human Actin, human Myosin, human Hemoglobin, human muscle creatine and human metallothionein.

Examples of polyadenylation signals useful to practice the present invention, especially in the production of a genetic vaccine for humans, include but are not limited to SV40 polyadenylation signals and LTR polyadenylation signals. In particular, the SV40 polyadenylation signal which is in pCEP4 plasmid (Invitrogen,  
5 San Diego Calif.), referred to as the SV40 polyadenylation signal, can be used.

In addition to the regulatory elements required for DNA expression, other elements may also be included in the DNA molecule. Such additional elements include enhancers. The enhancer may be selected from the group including but not limited to: human Actin, human Myosin, human Hemoglobin, human muscle  
10 creatine and viral enhancers such as those from CMV, RSV and EBV.

Genetic constructs can be provided with mammalian origin of replication in order to maintain the construct extrachromosomally and produce multiple copies of the construct in the cell. Plasmids pCEP4 and pREP4 from Invitrogen (San Diego, Calif.) contain the Epstein Barr virus origin of replication and nuclear antigen  
15 EBNA-1 coding region which produces high copy episomal replication without integration.

An additional element may be added which serves as a target for cell destruction if it is desirable to eliminate cells receiving the genetic construct for any reason. A herpes thymidine kinase (tk) gene in an expressible form can be included  
20 in the genetic construct. When the construct is introduced into the cell, tk will be produced. The drug gancyclovir can be administered to the individual and that drug will cause the selective killing of any cell producing tk. Thus, a system can be provided which allows for the selective destruction of vaccinated cells.

In order to be a functional genetic construct, the regulatory elements must be  
25 operably linked to the nucleotide sequence that encodes the target protein. Accordingly, it is necessary for the initiation and termination codons to be in frame with the coding sequence.

Open reading frames (ORFs) encoding the protein of interest and another or other proteins of interest may be introduced into the cell on the same vector or on  
30 different vectors. ORFs on a vector may be controlled by separate promoters or by a single promoter. In the latter arrangement, which gives rise to a polycistronic message, the ORFs will be separated by translational stop and start signals. The presence of an internal ribosome entry site (IRES) site between these ORFs permits the production of the expression product originating from the second ORF of

interest, or third, etc. by internal initiation of the translation of the bicistronic or polycistronic mRNA.

According to the invention, the genetic vaccine may be administered directly into the individual to be immunized or ex vivo into removed cells of the individual which are reimplanted after administration. By either route, the genetic material is introduced into cells which are present in the body of the individual. Routes of administration include, but are not limited to, intramuscular, intraperitoneal, intradermal, subcutaneous, intravenous, intraarterially, intraocularly and oral as well as transdermally or by inhalation or suppository. Preferred routes of administration include intramuscular, intraperitoneal, intradermal and subcutaneous injection. Genetic constructs may be administered by means including, but not limited to, traditional syringes, needleless injection devices, or microprojectile bombardment gene guns. Alternatively, the genetic vaccine may be introduced by various means into cells that are removed from the individual. Such means include, for example, ex vivo transfection, electroporation, microinjection and microprojectile bombardment. After the genetic construct is taken up by the cells, they are reimplanted into the individual. It is contemplated that otherwise non-immunogenic cells that have genetic constructs incorporated therein can be implanted into the individual even if the vaccinated cells were originally taken from another individual.

The genetic vaccines according to the present invention comprise about 1 nanogram to about 1000 micrograms of DNA. In some preferred embodiments, the vaccines contain about 10 nanograms to about 800 micrograms of DNA. In some preferred embodiments, the vaccines contain about 0.1 to about 500 micrograms of DNA. In some preferred embodiments, the vaccines contain about 1 to about 350 micrograms of DNA. In some preferred embodiments, the vaccines contain about 25 to about 250 micrograms of DNA. In some preferred embodiments, the vaccines contain about 100 micrograms DNA.

The genetic vaccines according to the present invention are formulated according to the mode of administration to be used. One having ordinary skill in the art can readily formulate a genetic vaccine that comprises a genetic construct. In cases where intramuscular injection is the chosen mode of administration, an isotonic formulation is preferably used. Generally, additives for isotonicity can include sodium chloride, dextrose, mannitol, sorbitol and lactose. In some cases, isotonic solutions such as phosphate buffered saline are preferred. Stabilizers include



gelatin and albumin. In some embodiments, a vaso-constriction agent is added to the formulation. The pharmaceutical preparations according to the present invention are provided sterile and pyrogen free.

5 Genetic constructs may optionally be formulated with one or more response enhancing agents such as: compounds which enhance transfection, i.e. transfecting agents; compounds which stimulate cell division, i.e. replication agents; compounds which stimulate immune cell migration to the site of administration, i.e. inflammatory agents; compounds which enhance an immune response, i.e. adjuvants or compounds having two or more of these activities.

10 In one embodiment, bupivacaine, a well known and commercially available pharmaceutical compound, is administered prior to, simultaneously with or subsequent to the genetic construct. Bupivacaine and the genetic construct may be formulated in the same composition. Bupivacaine is particularly useful as a cell stimulating agent in view of its many properties and activities when administered to  
15 tissue. Bupivacaine promotes and facilitates the uptake of genetic material by the cell. As such, it is a transfecting agent. Administration of genetic constructs in conjunction with bupivacaine facilitates entry of the genetic constructs into cells. Bupivacaine is believed to disrupt or otherwise render the cell membrane more permeable. Cell division and replication is stimulated by bupivacaine. Accordingly,  
20 bupivacaine acts as a replicating agent. Administration of bupivacaine also irritates and damages the tissue. As such, it acts as an inflammatory agent which elicits migration and chemotaxis of immune cells to the site of administration. In addition to the cells normally present at the site of administration, the cells of the immune system which migrate to the site in response to the inflammatory agent can come  
25 into contact with the administered genetic material and the bupivacaine. Bupivacaine, acting as a transfection agent, is available to promote uptake of genetic material by such cells of the immune system as well.

In addition to bupivacaine, mepivacaine, lidocaine, procains, carbocaine, methyl bupivacaine, and other similarly acting compounds may be used as response  
30 enhancing agents. Such agents acts a cell stimulating agents which promote the uptake of genetic constructs into the cell and stimulate cell replication as well as initiate an inflammatory response at the site of administration.

Other contemplated response enhancing agents which may function as transfecting agents and/or replicating agents and/or inflammatory agents and which  
35 may be administered include lectins, growth factors, cytokines and lymphokines

such as alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), gCSF, gMCSF, TNF, epidermal growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10 and IL-12 as well as collagenase, fibroblast growth factor, estrogen, dexamethasone, saponins, surface active agents such as immune-stimulating complexes (ISCOMS), Freund's incomplete adjuvant, LPS analog including monophosphoryl Lipid A (MPL), muramyl peptides, quinone analogs and vesicles such as squalene and squalane, hyaluronic acid and hyaluronidase may also be used administered in conjunction with the genetic construct. In some embodiments, combinations of these agents are co-administered in conjunction with the genetic construct. In other embodiments, genes encoding these agents are included in the same or different genetic construct(s) for co-expression of the agents.

With respect to HIV Env, Gag, and Pol nucleotide sequences of the invention, particularly through genetic immunization, may be used as therapeutics or prophylatics in the treatment of AIDS or HIV infection. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that includes the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the (e.g., the concentration of the test compound which achieves a half-maximal inhibition of viral infection relative to the amount of the event in the absence of the test compound) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography (HPLC).

The compounds (for genetic immunization) of the invention may, further, serve the role of a prophylactic vaccine, wherein the host produces antibodies and/or CTL responses against HIV Env, Gag, and Pol which responses then preferably serve to neutralize HIV viruses by, for example, inhibiting further HIV infection.

5 Administration of the compounds of the invention as a prophylactic vaccine, therefore, would comprise administering to a host a concentration of compounds effective in raising an immune response which is sufficient to elicit antibody and/or CTL responses to HIV Env, Gag, and Pol and/or neutralize HIV, by, for example, inhibiting HIV ability to infect cells. The exact concentration will depend upon the

10 specific compound to be administered, but may be determined by using standard techniques for assaying the development of an immune response which are well known to those of ordinary skill in the art.

#### *Env*

To improve the immune response to native gp160 and to expose the core protein for optimal antigen presentation and recognition, we have analyzed the

15 immune response to modified forms of the protein. The role of conserved N-linked glycosylation sites has been studied, and analogues of fusion intermediates have been developed. Expression vectors with deletions in the cleavage site (C), the fusion peptide (F), and the interspace (I) between the two heptad repeats were

20 termed  $\Delta$ CFI. Plasmid DNA vaccination has been a useful technology for the development and analysis of immunogens. This method of vaccination allows appropriate post-translational modification, proper intracellular trafficking, and antigen presentation. Direct injection of naked DNA either intramuscularly or intradermally in rodents induces immune responses, and the ability to easily modify

25 plasmid expression vectors to express different forms of HIV envelope proteins enables rapid and systematic testing of vaccine immunogens. In this disclosure, we have analyzed the immune response to modified Env candidates expressed in plasmids with modified codons to improve gene expression. Both antibody and CTL responses were analyzed after injection of plasmid DNA into muscle. A

30 modified gp140 DNA with improved ability to elicit antibody and CTL responses to HIV Env has now been identified that is envisioned as a prototype immunogen that can elicit broadly neutralizing antibody responses to HIV.

### Exposing the Core Protein of Viral Membrane Fusion Proteins

Described herein are modified HIV envelope proteins that improve the immune response to native gp160 and expose the core protein for optimal antigen presentation and recognition. Weissenhorn et al., *Molecular Cell*, 2, 605-616, 1998  
 5 proposes a core protein as a model for a fusion intermediate of viral glycoproteins, where the glycoproteins are characterized by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain direction and connects to an  $\alpha$  helix packed antiparallel to the core helices, as, for example, in the case of Ebola Zaire GP2, Murine Moloney Leukemia virus (MuMoLv) 55-residue segment  
 10 of the TM subunit (Mo-55), low-pH-treated influenza HA2, protease resistant core of HIV gp41, and SIV gp41 (Figure A). Thus, the strategy for improving the immune response by exposing the protease resistant core of HIV gp41 extends to other viral membrane fusion proteins that are characterized by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain  
 15 direction and connects to an  $\alpha$  helix packed antiparallel to the core helices.

The present approach involves a series of internal mutations designed to replace the cleavage site (C), the fusion domain (F), and the interspace (I) between the two heptad repeats all on a backbone of COOH-terminal truncations to expose the core protein of the viral membrane fusion protein Env, based on modified gp140  
 20 as a prototype immunogen. By replacement is meant deletions, insertions, and/or substitutions of amino acid residues. In one embodiment, deletions are meant (i.e., amino acids are deleted to create the  $\Delta$ CFI mutations).

In this embodiment, the  $\Delta$ C mutation is intended to eliminate proteolysis by deleting the gp120/gp41 cleavage site that links the envelope covalently to the  
 25 ectodomain by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%,  
 30 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the  $\Delta$ F mutation is intended to solubilize the molecule by deleting the fusion domain by 100%, 99%, 98%, 97%, 96%, 95%, 94%, 93%,



92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%,  
 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%,  
 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%,  
 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%,  
 5 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%,  
 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%,  
 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the  $\Delta I$  mutation is intended to stabilize oligomer  
 formation by deleting the interspace between the two heptad repeats by 100%, 99%,  
 10 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%,  
 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%,  
 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%,  
 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42,  
 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%,  
 15 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%,  
 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, the COOH-terminal truncation is intended to reduce  
 toxicity by deleting the cytoplasmic domain by 100%, 99%, 98%, 97%, 96%, 95%,  
 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%,  
 20 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%,  
 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%,  
 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42, 41%, 40%, 39%, 38%,  
 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%,  
 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%,  
 25 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.

In this embodiment, optionally, the COOH-terminal truncation is extended  
 so as to solubilize the molecule by deleting the transmembrane domain by 100%  
 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%,  
 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%,  
 30 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%,  
 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%,  
 43%, 42, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%,  
 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%,  
 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1%.



Amino acid substitutions may encompass those of a conserved or non-conserved nature. Presumably, a non-conserved substitution of a domain would act like a deletion of the domain. Conserved amino acid substitutions constitute switching one or more amino acids with amino acids of similar charge, size, and/or hydrophobicity characteristics. Non-conserved amino acid substitutions constitute switching one or more amino acids with amino acids of dissimilar charge, size, and/or hydrophobicity characteristics. The families of amino acids include the basic amino acids (lysine, arginine, histidine); the acidic amino acids (aspartic acid, glutamic acid); the non-polar amino acids (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan); the uncharged polar amino acids (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine); and the aromatic amino acids (phenylalanine, tryptophan, and tyrosine). One or more substitutions may be introduced to achieve the  $\Delta C$  mutation intended to eliminate proteolysis by acting like a deletion of the gp120/gp41 cleavage site to link the envelope covalently to the ectodomain, the  $\Delta F$  mutation intended to solubilize the molecule by acting like a deletion of the fusion domain, the  $\Delta I$  mutation intended to stabilize oligomer formation by acting like a deletion of the interspace between the two heptad repeats, the COOH-terminal truncation intended to reduce toxicity by acting like a deletion of the cytoplasmic domain, and, optionally, the COOH-terminal truncation extended so as to solubilize the molecule by acting like a deletion of the transmembrane domain.

Amino acid insertions may constitute single amino acid residues or stretches of residues. The insertions may be made at the carboxy or amino terminal end of a domain, as well as at a position internal to the domain. Such insertions will generally range from 2 to 15 amino acids in length. One or more insertions may be introduced to achieve the  $\Delta C$  mutation intended to eliminate proteolysis by acting like a deletion of the gp120/gp41 cleavage site to link the envelope covalently to the ectodomain, the  $\Delta F$  mutation intended to solubilize the molecule by acting like a deletion of the fusion domain, the  $\Delta I$  mutation intended to stabilize oligomer formation by acting like a deletion the interspace between the two heptad repeats, the COOH-terminal truncation intended to reduce toxicity by acting like a deletion of the cytoplasmic domain, and, optionally, the COOH-terminal truncation extended so as to solubilize the molecule by acting like a deletion of the transmembrane domain.

The nucleic acids of the present invention are optionally DNA, RNA, or mRNA. Most typically, the nucleic acids are provided by recombinantly making a DNA, which is expressed in a cell as RNA and/or as mRNA. Given the strategy for making the nucleic acids of the present invention, one of skill can construct a variety of clones containing functionally equivalent nucleic acids. Cloning methodologies to accomplish these ends, and sequencing methods to verify the sequence of nucleic acids are well known in the art. Examples of appropriate cloning and sequencing techniques, and instructions sufficient to direct persons of skill through many cloning exercises are found in Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volume 152, Academic Press, Inc., San Diego, Calif. (Berger); Sambrook, et al. (1989) Molecular Cloning--A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, N.Y., (Sambrook); and Current Protocols in Molecular Biology, F. M. Ausubel, et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1994 Supplement) (Ausubel). Product information from manufacturers of biological reagents and experimental equipment also provide information useful in known biological methods. Such manufacturers include the SIGMA chemical company (Saint Louis, Mo.), R&D systems (Minneapolis, Minn.), Pharmacia LKB Biotechnology (Piscataway, N.J.), CLONTECH Laboratories, Inc. (Palo Alto, Calif.), Chem Genes Corp., Aldrich Chemical Company (Milwaukee, Wis.), Glen Research, Inc., GIBCO BRL Life Technologies, Inc. (Gaithersburg, Md.), Fluka Chemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Switzerland), Invitrogen, San Diego, Calif., and Applied Biosystems (Foster City, Calif.), as well as many other commercial sources known to one of skill.

The nucleic acid compositions of this invention, whether RNA, cDNA, mRNA, genomic DNA, or a hybrid of the various combinations, are isolated from biological sources or synthesized in vitro. The nucleic acids of the present invention are present in transformed or transfected whole cells, in transformed or transfected cell lysates, or in a partially purified or substantially pure form.

In vitro amplification techniques suitable for amplifying sequences to provide a nucleic acid or for subsequent analysis, sequencing or subcloning are known. Examples of techniques sufficient to direct persons of skill through such in vitro amplification methods, including the polymerase chain reaction (PCR) the

ligase chain reaction (LCR), Q $\beta$ -replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as Mullis, et al., (1987) U.S. Pat. No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis, et al. eds) Academic Press Inc. San Diego, Calif. (1990) (Innis); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3:81-94; Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86:1173 (1989); Guatelli, et al., Proc. Natl. Acad. Sci. USA, 87:1874 (1990); Lomell, et al., J. Clin. Chem., 35:1826 (1989); Landegren, et al., Science, 241:1077-1080 (1988); Van Brunt, Biotechnology, 8:291-294 (1990); Wu and Wallace, Gene, 4:560 (1989); Barringer, et al., Gene, 89:117 (1990), and Sooknanan and Malek, Biotechnology, 13:563-564 (1995). Improved methods of cloning in vitro amplified nucleic acids are described in Wallace, et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids (up to 40 kb) are summarized in Cheng, et al., Nature, 369:684-685 (1994) and the references therein. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, Ausubel, Sambrook, Innis, and Berger, all supra.

One of skill will recognize many ways of generating alterations in a given nucleic acid construct. Such well-known methods include site-directed mutagenesis, PCR amplification using degenerate oligonucleotides, exposure of cells containing the nucleic acid to mutagenic agents or radiation, chemical synthesis of a desired oligonucleotide (e.g., in conjunction with ligation and/or cloning to generate large nucleic acids) and other well-known techniques. See, Giliman and Smith, Gene 8:81-97 (1979), Roberts, et al., Nature, 328:731-734 (1987) and Sambrook, Innis, Ausubel, Berger, and Mullis (all supra).

Most modifications to nucleic acids are evaluated by routine screening techniques in suitable assays for the desired characteristic. For instance, changes in the immunological character of encoded polypeptides can be detected by an appropriate immunological assay. For instance, changes in the cellular immunological character of the polypeptide can be detected by an appropriate antibody or CTL assay. Modifications of other properties such as nucleic acid hybridization to a complementary nucleic acid, redox or thermal stability of encoded proteins, hydrophobicity, susceptibility to proteolysis, or the tendency to aggregate are all assayed according to standard techniques.

A wide variety of formats and labels are available and appropriate for detection of polypeptide sequences. These include analytic biochemical methods such as spectrophotometry, radiography, electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, and various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), western blot assays, immunofluorescent assays, and the like. Several commercially available ELISA assays for the detection of retroviral components, including Env domains, are available, allowing one of skill to detect Env in biological samples.

Similarly, the detection of the nucleic acids of the present invention proceeds by well known methods such as Southern analysis, northern analysis, gel electrophoresis, PCR, radiolabeling and scintillation counting, and affinity chromatography. Many assay formats are appropriate, including those reviewed in Tijssen (1993) *Laboratory Techniques in biochemistry and molecular biology--hybridization with nucleic acid probes parts I and II*, Elsevier, New York and Choo (ed) (1994) *Methods In Molecular Biology Volume 33--In Situ Hybridization Protocols*, Humana Press Inc., New Jersey (see also, other books in the *Methods in Molecular Biology* series); see especially, Chapter 21 of Choo (id.) "Detection of Virus Nucleic Acids by Radioactive and Nonisotopic in Situ Hybridization". Finally, PCR is also routinely used to detect nucleic acids in biological samples (see, Innis, *supra*, for a general description of PCR techniques).

In one preferred embodiment, antibodies are used to detect polypeptide sequences. Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many anti-HIV antibodies are available. See, e.g., Coligan (1991) *Current Protocols in Immunology*, Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY; Stites, et al. (eds.) *Basic and Clinical Immunology* (4th ed.), Lange Medical Publications, Los Altos, Calif., and references cited therein; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.), Academic Press, New York, N.Y.; and Kohler and Milstein, *Nature*, 256:495-497 (1975). Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. See, Huse, et al., *Science*, 246:1275-1281 (1989); and Ward, et al., *Nature*, 341:544-546 (1989). Specific monoclonal and

polyclonal antibodies and antisera will usually bind with a KD of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most typically and preferably, 0.01  $\mu$ M or better.

#### Development of HIV Env Vectors

5           To develop Env glycoprotein variants that might effectively induce humoral and cellular immunity, a series of plasmid expression vectors were generated (Fig. I). HIV Env is encoded by nucleic acid sequences that contain RNA structures that limit gene expression (Andre S et al, 1998, J Virol, 72:1497-503; Qiu JT et al., 1999, J Virol, 73:9145-52; Rana TM et al., 1999, Arch Biochem Biophys, 365:175-85; 10 Roebuck KA, Saifuddin M, 1999, Gene Expr, 8:67-84; Romano G, et al., 1999, J Cell Biochem, 75:357-68; Schneider R et al., 1997, J Virol, 71:4892-903). These vectors were therefore synthesized using codons found in human genes that allow these structures to be eliminated without affecting the amino acid sequence. Full-length HIV Env (gp160) was highly expressed in the absence of HIV accessory 15 proteins at levels  $\geq 10$ -fold higher than Rev-dependent viral gp 160 in transfected 293 cell by Western blot analysis (Fig. IIA), and the relevant mutant proteins were detected at the expected apparent molecular weights (Fig. IIC-D). As might be anticipated, gp160 expressed from the synthetic gene was not efficiently processed in transfected 293 cells, presumably because over-expression of gp160 saturates the 20 cellular proteases responsible for cleavage. (Binley JM et al., 2000, J. Virol, 74:627-643.) Synthetic HIV gp160 induced toxicity in transfected cells, with cell rounding and detachment evident within 48 hours (Fig. IIIA vs B). This cytotoxicity was reduced by elimination of the COOH-terminal cytoplasmic domain. Env protein that terminated at amino acid 752 (gp150) was less cytotoxic than gp160, while the 25 shorter proteins (gp145 and gp140) produced little or no effect (Fig. IIIC, D, and E, respectively).

          To alter Env immunogenicity, two different approaches were explored. First the effects of glycosylation on cellular and humoral immunity were evaluated by analysis of mutants in which conserved N-linked glycosylation sites were eliminated 30 by site-directed mutagenesis. Two sets of mutations were introduced into both gp160 and gp150 (Fig. I). The first set included eleven potential sites ( $\Delta$ gly11), and the second set included an additional 6 sites downstream ( $\Delta$ gly17). Expression studies showed that the glycosylation mutants were efficiently expressed, and the



glycosylation mutant protein was appropriately reduced in size compared to wild type gp160 or gp150, consistent with reduced N-linked glycosylation (Fig IIC).

5 The second approach involved a series of internal deletions designed to stabilize and expose functional domains of the protein that might be present in an extended helical structure prior to the formation of the six-member coiled-coil structure in the hairpin intermediate (Chan DC et al., 1997, Cell, 89:263-73; Weissenhorn W et al., 1997, Nature, 387:426-30). To generate this putative pre-hairpin structure, the cleavage site was removed to prevent the proteolytic processing of the envelope and stabilize the protein by linking it covalently to the gp41 extracellular and/or transmembrane domain. To reduce toxicity and enhance stability, the fusion peptide domain was deleted. The heptad repeats in the envelope protein are important tertiary structure domains involved in the ability of the envelope protein to form trimers (Lu M, Kim PS, 1997, J Biomol Struct Dyn, 15:465-71). The sequence between the heptad repeats was removed to stabilize the formation of trimers and eliminate formation of the hairpin intermediate. These  $\Delta$ CFI deletions were introduced into full-length gp160 and COOH-terminal truncation mutants. Though cells transfected with vectors encoding gp140 $\Delta$ CFI, gp145 $\Delta$ CFI and gp160 readily expressed these proteins (Fig. IID), only gp140 $\Delta$ CFI, which lacks the transmembrane domain, was readily detected in the supernatant (Fig. IV), indicating that it can give rise to soluble antigen.

#### Immunogenicity of Env Mutants After DNA Vaccination

25 The ability of these Env proteins to elicit an immune response was determined in mice by injection with these plasmid DNA expression vectors. Antibody responses were monitored by the ability of antisera from injected mice to immunoprecipitate wild type gp160 from cell lysates by SDS-PAGE and Western blotting (Fig. V). In some cases, antibody reactivity was also confirmed by immunofluorescence. To quantitate the antibody response, immunoprecipitation followed by Western blotting with different dilutions of immunized mouse sera was tested. The intensity of the gp160 band was determined by densitometry and standardized relative to a positive control sera used to normalize data between experiments. The approximately linear dose response of gp160 intensity with serum dilution allowed quantification of the anti-gp160 antibody response in mouse sera. Data from immunized mice showed that none of the wild type Env proteins, neither the gp160, gp150, gp145, nor gp140 COOH-terminal truncations generated

consistently high antibody responses (Fig. VA). Of these proteins, gp140 was somewhat more effective than gp160 in generating antibody responses but remained low and inconsistent. Immunization with vectors designed to express glycosylation deficient envelope proteins also did not improve the humoral response. In contrast, the  $\Delta$ CFI mutants in the Env truncation vectors substantially increased anti-gp160 antibody response (Fig. VA; gp145, gp140, and gp128). The longer  $\Delta$ CFI Env proteins (gp160 and gp150) did not show comparable enhancement of this response. The gp140 ( $\Delta$ CFI) provided more consistent and a greater increase in antibody response than gp128 $\Delta$ CFI (Fig VB). In all cases, these vectors that encoded gp140 $\Delta$ CFI, but not wild type gp140, induced antibodies that were reactive to native gp160 (Fig. VC).

To determine whether these modifications of Env adversely affected CTL responses, spleen cells from immunized mice were tested for their ability to lyse relevant target cells. All mice immunized with codon-altered Env vectors, including COOH-terminal deletion mutations (Fig. VIA) or glycosylation mutants (Fig. VIB), elicited strong CTL responses directed to cell lines pulsed with HIV Env peptides. These findings were also confirmed using stably transfected cells that expressed Env. Importantly, gp140 ( $\Delta$ CFI), which elicited increased antibody responses relative to the comparable wild type Env, readily induced CTL responses to native Env, as did other  $\Delta$ CFI mutants (Fig. VIC). Addition of anti-CD8 antibody inhibited cytolytic activity, as did depletion using magnetic beads coupled with anti-CD8 antibody, thus confirming a cytotoxic T cell response to these immunogens after genetic immunization (Fig. VID). This response was detectable for at least six months after immunization.

#### 25 Glycosylation and $\Delta$ CFI HIV Env Mutants

To develop DNA vaccine candidates for HIV, we developed a series of synthetic genes designed to express HIV Env mutants in human cells. In the absence of HIV regulatory proteins, these codon-altered envelope protein genes expressed well in human cells. Like other DNA vaccines, immunization with these vectors elicited strong CTL responses in mice, and antibody responses were not robust in mice immunized with wild type Env expression vectors. Mutations in highly conserved N-linked glycosylation sites did not significantly alter humoral or cellular immune response to native Env. In contrast, a mutant Env with deletions in the cleavage site, fusion domain, and a region between the heptad repeats elicited a

more potent humoral immune response and retained its ability to stimulate Env-specific CTL.

Recent reports suggest that gp160 forms trimers *in vivo* and the domain required for trimer formation resides in the ectodomain of the gp41 (Yang X et al., 2000, J Virol, 74:5716-25). Such trimeric forms of HIV envelope protein are likely to present different epitopes to the immune system compared to monomeric gp120. In addition to the linear epitopes in the envelope, this trimeric structure is likely to expose conformational epitopes important for B cell triggering of a relevant antibody response. In this regard, gp140 ( $\Delta$ CFI), which induced the greatest antibody response, is released in a soluble form (Fig. IVB). In contrast, wild type Env did not elicit high titer antibody responses. The toxicity of Env in mammalian cells has been seen and could limit both the amount and duration of envelope protein expression *in vivo* that would affect immunogenicity. The envelope is also heavily glycosylated, and removal of partial or complete gp120 glycosylation sites has resulted in higher titers of strain-specific neutralizing antibody responses to mutant SIVs in monkeys (Binley JM et al., 1998, AIDS Res Hum Retroviruses, 14 :191-8; Reitter JN et al., 1998, Nat Med, 4:679-84). Though it seemed reasonable that deglycosylation would reveal epitopes otherwise masked in the native protein, we did not observe enhanced immune reactivity by DNA vaccination using different glycosylation site mutants, both in gp160 and gp150. This difference with the previous study is likely due to the fact that DNA vaccination rather than viral infection was utilized for immunization. Though glycosylation mutants are unlikely to prove helpful with this former method of immunization, we envision that modification of glycosylation sites will be effective with other vectors or adjuvants.

HIV-1 Env is proteolytically cleaved by a cellular convertase into gp120 and gp41 (Eckert DM et al., 1999, Cell, 99:103-15). The gp41 subunit is composed of cytoplasmic, transmembrane, and ectodomain segments. The role of the ectodomain of the envelope in membrane fusion, particularly its hydrophobic glycine-rich fusion peptide, is well established. Two regions with heptad coiled-coil repeats in the ectodomain of gp41 are involved in viral fusion (Eckert DM et al., 1999, Cell, 99:103-15; Weissenhorn W et al., 1996, EMBO J, 15:1507-14). Upon fusion, these two alpha helices, connected via a disulfide-stabilized loop (Gallagher WR et al., 1989, AIDS Res Hum Retroviruses, 5:431-40; Kent KA, Robinson J, 1996, AIDS, 10 Suppl A:S107-S114; Sattentau QJ et al., 1995, Virology, 206:713-7), presumably

undergo a transient conformational change to a fusion active state. These changes allow the formation of a six-member helical hairpin intermediate structure that presumably exposes the fusion peptide at the NH<sub>2</sub>-terminus of gp41, allowing fusion to the target cell membrane (Binley J, Moore JP, [published erratum appears in Nature 1997 Sep 11;389(6647):131], 1997, Nature, 387 :346-8; LaCasse R et al., 1999, Science, 283:357-62). The ΔCFI mutation was intended to eliminate cleavage of gp140, remove the unstable hydrophobic region and stabilize oligomer formation. Though detailed structural data is not yet available on this protein, these mutations apparently stabilize Env in a conformation that elicits both humoral and cellular immune responses. For example, the neutralizing epitope in the ectodomain of gp41 (Muster T et al., 1994, J Virol, 68:4031-4) is present in the series of deletions and truncations of the envelope and gp140 (ΔCFI) is reactive with the 2F5 neutralizing monoclonal antibody that binds to this epitope. Importantly, these immunogens also induced CTL responses to Env. Though gp128 (ΔCFI) induced slightly more potent CTL activity, gp140 (ΔCFI) was better able to elicit such responses, both to peptide-pulsed cells and stably transduced target cells. Thus the enhanced humoral immune response introduced by this vaccine candidate did not appear to diminish the CTL response. Taken together, these results indicate that gp140ΔCFI serves as an improved immunogen that can more effectively elicit an antibody response against the envelope by DNA vaccination while preserving its ability to induce a CTL response.

### *Gag and Pol*

In this disclosure, we have prepared synthetic HIV-1 B clade Gag and Pol expression vectors that are based on human (h) codon usage. These vectors encode hGag-Pol and its derivatives, hGag, hPol and an hGag-Pol fusion protein. The synthetic Gag-Pol genes show little nucleotide homology to HIV-1 but are the same in protein sequence. The modified Gag-Pol genes were subcloned into a eukaryotic plasmid expression vector for expression and DNA immunization studies. Synthetic Gag-Pol genes allowed high level Rev-independent expression of HIV-1 Gag-Pol precursor proteins in human and mouse cell lines and induced significant cellular and humoral responses in mice. The Gag-Pol fusion protein induced the broadest responses to Gag and Pol determinants and thus is envisioned as a prototype immunogen that maximizes epitope presentation.

### Eliminating The Frame Shift Site To Create Viral Polyproteins

Described herein are HIV Gag-Pol fusion proteins encoded by a continuous open reading frame so to improve the immune response to native Gag and Pol. In some viruses, translational frame shifting is exploited during protein synthesis. Specific sequences in the RNA are required for the frame shifting. The viral RNA sequences cause ribosomal slippage so that viral proteins are produced in non-equivalent ratios. For example, during translation in HIV, the ribosomes shift reading frames to synthesize Gag precursor protein and the gag-pol fusion protein in a 20:1 ratio. The strategy here is to maximize epitope presentation by transcribing an immunogen from a continuous open reading frame by eliminating the frame shift site. Thus, the strategy for improving the immune response by the use of a HIV Gag-Pol fusion protein encoded by a continuous open reading frame extends to other viral proteins that are produced in non-equivalent ratios by virtue of translational frame shifting.

The present invention involves HIV Gag-Pol fusion proteins encoded by a single continuous open reading frame due to mutation of the frame shift site. The frame shift site is mutated by deletions, insertions, and/or substitutions of nucleotides to create a single continuous open reading frame. In one embodiment, deletions are meant (i.e., nucleotides are deleted to create the same open reading frame).

The frame shift site is a mutated frame shift to create a single continuous open reading frame. For example, a set of similar retroviral gag-pol frame shift sites are optionally made for a given fusion protein, for example, by synthesizing different gag-pol frame shift regions and cloning the sequences appropriately, or by site-directed mutagenesis of a given frame shift clone. The efficacy of the frame shift sites are assessed by measuring the production of the fusion protein. The sequence that shows the highest level of expression is a "optimized" frame shift mutation for the set assessed. Alternatively, where a particular level of expression is desired, a frame shift site from a particular set of possible frame shift sites which is closest to the desired activity level is considered to be "optimized."

Although a full length Gag sequence is preferred for use in the fusion protein of the present invention, Gag is optionally deleted of subsequences without negating a polyepitope response. For example, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be deleted while preserving the polyepitope



response. Alternatively, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be substituted while preserving the polyepitope response. Alternatively, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be interrupted by insertions while preserving the polyepitope response. Optionally, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of p2, regions of the nucleocapsid protein (p7), regions of p1, and regions of p6 can be mutated to inactivate these proteins.

Likewise, a full length Pol sequence is preferred for use in the fusion protein of the present invention, and Pol is optionally deleted of subsequences without negating a polyepitope response. For example, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be deleted while preserving the polyepitope response. Alternatively, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be substituted while preserving the polyepitope response. Alternatively, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be interrupted by insertions while preserving the polyepitope response. Optionally, regions of the protease protein, regions of the reverse transcriptase protein, and regions of the integrase protein can be mutated to inactivate these enzymes.

In other embodiments, the Gag-Pol fusion proteins are co-expressed with other proteins, either as fusion proteins or separately, preferably with Env sequence proteins, such as the modified HIV Env  $\Delta$ CFI proteins described herein.

In another aspect, the invention involves chimeric nucleic acid molecules. The chimeric nucleic acid molecules of the invention typically have a retroviral gag nucleic acid sequence and a retroviral pol nucleic acid sequence in the same open reading frame due to mutation of the frame shift site. The continuous open reading frame encodes a fusion protein such as those described above.

The chimeric nucleic acids of the present invention are optionally DNA, RNA, or mRNA. Most typically, the chimeric nucleic acids are provided by recombinantly making a DNA, which is expressed in a cell as RNA and/or as mRNA. Given the strategy for making the chimeric nucleic acids of the present invention, one of skill can construct a variety of clones containing functionally

equivalent nucleic acids. Cloning methodologies to accomplish these ends, and sequencing methods to verify the sequence of nucleic acids are well known in the art. Examples of appropriate cloning and sequencing techniques, and instructions sufficient to direct persons of skill through many cloning exercises are found in

5 Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volume 152, Academic Press, Inc., San Diego, Calif. (Berger); Sambrook, et al. (1989) Molecular Cloning--A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, N.Y., (Sambrook); and Current Protocols in Molecular Biology, F. M. Ausubel, et al., eds., Current

10 Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1994 Supplement) (Ausubel). Product information from manufacturers of biological reagents and experimental equipment also provide information useful in known biological methods. Such manufacturers include the SIGMA chemical company (Saint Louis, Mo.), R&D systems (Minneapolis, Minn.),

15 Pharmacia LKB Biotechnology (Piscataway, N.J.), CLONTECH Laboratories, Inc. (Palo Alto, Calif.), Chem Genes Corp., Aldrich Chemical Company (Milwaukee, Wis.), Glen Research, Inc., GIBCO BRL Life Technologies, Inc. (Gaithersburg, Md.), Fluka Chemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Switzerland), Invitrogen, San Diego, Calif., and Applied Biosystems (Foster City, Calif.), as well as many other commercial sources known to one of skill.

20

The chimeric nucleic acid compositions of this invention, whether RNA, cDNA, mRNA, genomic DNA, or a hybrid of the various combinations, are isolated from biological sources or synthesized in vitro. The chimeric nucleic acids of the present invention are present in transformed or transfected whole cells, in

25 transformed or transfected cell lysates, or in a partially purified or substantially pure form.

In vitro amplification techniques suitable for amplifying sequences to provide a nucleic acid or for subsequent analysis, sequencing or subcloning are known. Examples of techniques sufficient to direct persons of skill through such in

30 vitro amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Q $\beta$ -replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as Mullis, et al., (1987) U.S. Pat. No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis, et al. eds) Academic Press Inc. San Diego, Calif.

(1990) (Innis); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3:81-94; Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86:1173 (1989); Guatelli, et al., Proc. Natl. Acad. Sci. USA, 87:1874 (1990); Lomell, et al., J. Clin. Chem., 35:1826 (1989); Landegren, et al., Science, 241:1077-1080 (1988);  
5 Van Brunt, Biotechnology, 8:291-294 (1990); Wu and Wallace, Gene, 4:560 (1989); Barringer, et al., Gene, 89:117 (1990), and Sooknanan and Malek, Biotechnology, 13:563-564 (1995). Improved methods of cloning in vitro amplified nucleic acids are described in Wallace, et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids (up to 40 kb) are summarized in Cheng, et al., Nature,  
10 369:684-685 (1994) and the references therein. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, Ausubel, Sambrook, Innis, and Berger, all supra.

One of skill will recognize many ways of generating alterations in a given  
15 nucleic acid construct. Such well-known methods include site-directed mutagenesis, PCR amplification using degenerate oligonucleotides, exposure of cells containing the nucleic acid to mutagenic agents or radiation, chemical synthesis of a desired oligonucleotide (e.g., in conjunction with ligation and/or cloning to generate large nucleic acids) and other well-known techniques. See, Giliman and Smith, Gene  
20 8:81-97 (1979), Roberts, et al., Nature, 328:731-734 (1987) and Sambrook, Innis, Ausubel, Berger, and Mullis (all supra).

Most modifications to nucleic acids are evaluated by routine screening techniques in suitable assays for the desired characteristic. For instance, changes in the immunological character of encoded polypeptides can be detected by an  
25 appropriate immunological assay. For instance, changes in the cellular immunological character of the polypeptide can be detected by an appropriate antibody or CTL assay. Modifications of other properties such as nucleic acid hybridization to a complementary nucleic acid, redox or thermal stability of encoded proteins, hydrophobicity, susceptibility to proteolysis, or the tendency to aggregate  
30 are all assayed according to standard techniques.

A wide variety of formats and labels are available and appropriate for detection of fusion protein sequences. These include analytic biochemical methods such as spectrophotometry, radiography, electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography  
35 (TLC), hyperdiffusion chromatography, and the like, and various immunological

methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), western blot assays, immunofluorescent assays, and the like. Several commercially available ELISA assays for the detection of retroviral components, including Env domains, are available, allowing one of skill to detect Env in biological samples.

Similarly, the detection of the chimeric nucleic acids of the present invention proceeds by well known methods such as Southern analysis, northern analysis, gel electrophoresis, PCR, radiolabeling and scintillation counting, and affinity chromatography. Many assay formats are appropriate, including those reviewed in Tijssen (1993) *Laboratory Techniques in biochemistry and molecular biology--hybridization with nucleic acid probes parts I and II*, Elsevier, New York and Choo (ed) (1994) *Methods In Molecular Biology Volume 33--In Situ Hybridization Protocols*, Humana Press Inc., New Jersey (see also, other books in the *Methods in Molecular Biology* series); see especially, Chapter 21 of Choo (id.) "Detection of Virus Nucleic Acids by Radioactive and Nonisotopic in Situ Hybridization". Finally, PCR is also routinely used to detect nucleic acids in biological samples (see, Innis, *supra*, for a general description of PCR techniques).

In one preferred embodiment, antibodies are used to detect polypeptide sequences. Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many anti-HIV antibodies are available. See, e.g., Coligan (1991) *Current Protocols in Immunology*, Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY; Stites, et al. (eds.) *Basic and Clinical Immunology* (4th ed.), Lange Medical Publications, Los Altos, Calif., and references cited therein; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.), Academic Press, New York, N.Y.; and Kohler and Milstein, *Nature*, 256:495-497 (1975). Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. See, Huse, et al., *Science*, 246:1275-1281 (1989); and Ward, et al., *Nature*, 341:544-546 (1989). Specific monoclonal and polyclonal antibodies and antisera will usually bind with a KD of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most typically and preferably, 0.01  $\mu$ M or better.

### Expression of Synthetic HIV-1 Gag and Pol Genes

Four synthetic HIV-1 Gag-and/or Pol expression vectors, hGag-Pol, hGag-Pol $\Delta$ Fs $\Delta$ Pr, hPol and hGag genes were prepared (Fig. VII). To confirm expression, the synthetic or viral Gag-Pol genes were transiently transfected into 293T cells, a human kidney-derived cell line. When cell lysates were analyzed by immunoblotting with human anti-HIV-1 IgG (Fig. VIIIA), monoclonal anti-p24 (Fig. VIIIB), and rabbit anti-RT (Fig. VIIIC), Gag p55, Pol p110 and Gag-Pol p160 precursor proteins were detected in hGag, hPol, and hGag-Pol fusion plasmids transfected 293T cells, as was expected. Mature virion proteins, p24 and RTp66, were detected in the hGag-Pol gene transfected cells (Figs. VIIIA, B and C). This might be a result of the activation of protease inside cells which was itself a result of the high-level expression of Gag and Gag-Pol protein (Karacostas, V et al., 1993, Virology, 193:661-671). The expression of Gag precursor proteins from codon-altered vectors was  $\geq 10$ -fold higher than viral Gag-Pol (Fig. VIII), determined by quantitative phosphorimaging. The level of accumulated Gag-Pol fusion protein was 100-fold higher in cells transfected with hGag-Pol compared to viral Gag-Pol. Virus-like particles were released from the hGag gene transfected cells (Fig. IX), detected by transmission electron microscopy. Though such particles were observed at a lower frequency with hGag-Pol, no particles were seen in cells transfected with hGag-Pol $\Delta$ Fs $\Delta$ Pr or hPol vectors. Stable expression of HIV-1 Gag and Pol proteins from codon-optimized genes in mouse CT26 and BH10ME cells was also observed (Fig. VIID).

### Induction of HIV-1 Gag and Pol CTL Responses in Mice by DNA Vaccination

To evaluate the cellular immune response to HIV-1 Gag and Pol proteins, Balb/C female mice were injected intramuscularly with the eukaryotic expression vector plasmids containing the codon-optimized genes. Two weeks after the final vaccination, splenocytes were harvested from the immunized mice and sensitized with either Gag or Pol peptide-pulsed naïve mouse splenocytes. One week later, CTL responses were analyzed using a 5-hour chromium release assay.

CTL responses specific to HIV-1 Gag and/or Pol were first analyzed using Gag or Pol peptide-pulsed BC10ME cells, or mouse fibrosarcoma cell lines derived from B/C-N cells (Collins, JL et al., 1981, J Exp. Med., 153:89-106). Immunization with hGag, hGag-Pol $\Delta$ FS $\Delta$ Pr or hGag-Pol genes induced comparably strong CTL responses specific to Gag (Fig. XA); however, after immunization with hPol, hGag-



Pol $\Delta$ FS $\Delta$ Pr or hGag-Pol genes, only the fusion protein, hGag-Pol $\Delta$ FS $\Delta$ Pr, and hPol to a lesser extent, elicited a marked CTL response to Pol (Fig. XB). To confirm that the specific killing in the CTL assays was induced by CD8<sup>+</sup> cytotoxic T lymphocytes, CD4<sup>+</sup> or CD8<sup>+</sup> cells were depleted from sensitized splenocytes by Dynal beads (Dynal, Inc., Lake Success, NY). Depletion of CD8<sup>+</sup> cells abolished the specific lysis in the hGag-Pol  $\Delta$ FS $\Delta$ Pr gene-immunized mice, while depletion of CD4<sup>+</sup> had little effect on lysis (Fig. XC), suggesting that CD8<sup>+</sup> lymphocytes were responsible for specific cytotoxicity.

The responses were further analyzed and confirmed with the hGag or hPol gene transduced syngeneic CT26 and BC10ME cell lines. Responses to Gag in the mice immunized with the hGag, hGag-Pol $\Delta$ FS $\Delta$ Pr or hGag-Pol genes were similar when peptide-pulsed cells were used as targets in the CTL assay (Fig. XIA). Mice immunized with the hPol gene generated a specific response to HIV-1 Pol on BC10ME cell lines stably expressing Pol as target cells (Fig. XIB). The same results have been observed with CT26 cell lines. These stably transfected cell lines were therefore more sensitive as target cells than peptide-pulsed cells in the Pol CTL assays.

#### Antibody Response in the Immunized Mice

Sera from mice immunized with different plasmids was analyzed with a p24 ELISA. hGag immunized mice demonstrated the highest p24 antibody titers (Fig. XIA). Unexpectedly, hGag-Pol virus-like particles elicited the lowest levels of p24 antibody. Similar results were observed by Western blotting with pooled sera (Fig. XIIB). The HIV-1 Pol specific antibodies were not detected by a commercially available Western blotting kit (Fig. XIIB), but antibodies to Pol were detected in mice immunized with hPol and hGag-Pol $\Delta$ FS $\Delta$ Pr with a more sensitive method, IP/Western blotting (Fig. XIIC). Presumably, this assay is more sensitive and better able to detect native conformational epitopes. Though such antibodies were found in mice immunized with Pol and Gag-Pol fusion proteins in this assay, minimal response was detected in the mice immunized with hGag-Pol. Though both immunogens elicited similar Gag responses, the Gag-Pol fusion protein was therefore more effective in the stimulation of CTL and antibody responses to Pol.

#### HIV Gag-Pol Fusion Proteins

In this disclosure, HIV-1 B-clade Gag and Pol genes were modified to increase Rev-independent expression of HIV-1 Gag-Pol proteins. This modification

allowed synthesis of HIV Gag and Pol, as well as fusion proteins at levels 10- to 100-fold higher than the corresponding Rev-dependent viral gene in the absence of Rev and RRE elements. These viral proteins were recognized by standard polyclonal and monoclonal antibodies (Fig. VIII), and immature VLP were produced and released from 293T cells transfected with hGag (Fig. IX).

The immune response induced by Rev-independent Gag-Pol was directed against both Gag and Pol determinants. The hGag, hGag-Pol fusion and hGag-Pol all induced strong CTL responses specific for Gag in mice immunized with plasmid DNA, but a significant Pol response was elicited only in the mice immunized with the hGag-Pol $\Delta$ FS $\Delta$ Pr or Pol alone. Because immunization with hGag-Pol gene failed to induce detectable cellular or humoral responses to HIV-1 Pol protein, these findings indicate that the Gag-Pol fusion protein induces a broader range of responses and allows delivery of an immunogen with a larger number of epitopes in a single continuous open reading frame. During viral replication, viral gag-pol produces Gag precursor protein and the gag-pol fusion protein by frame shifting in a 20:1 ratio (Wilson, W et al., 1988, Cell, 55:1159-1169). The deletion of a frame shift site in hGag-Pol $\Delta$ FS $\Delta$ Pr results in production of only the Gag-Pol fusion protein. Expression of Gag-Pol proteins alone in human cells is not adequate to form releasable viral particles because HIV-1 viral assembly requires Gag precursor proteins (Park, J and CD Morrow, 1992, J Virol, 66:6304-6313; Smith, AJ et al., 1993, J Virol, 67:2266-2275). The ability of hGag-Pol $\Delta$ FS $\Delta$ Pr to elicit strong Gag and Pol specific CTL responses in mice may be explained by high level expression of Gag-Pol fusion protein and its retention on the inside of cells, which could create a new condition not existing during normal viral replication and could provide sufficient proteins for antigen presentation. Moreover, the mutation in viral protease prevents viral protein from intracellular activation and reduced cellular toxicity. Overexpression of this polyprotein is also likely to affect its intracellular localization/transport and is envisioned as improving antigen presentation.

The Pol gene of HIV-1 is the precursor protein for viral protease, reverse transcriptase and integrase, which are crucial to viral replication (Kohl, NE et al., 1988, Proc. Natl. Acad. Sci. USA, 85:4686-4690.). Retroviral extracellular maturation, resulting from self-activation of protease after release produces mature RT and IN which have important functions in reverse transcription and integration respectively, for HIV-1, HIV-2, and SIV replication. The catalytic cores of these

enzymes have relatively conserved domains in order to preserve their functions and thus are envisioned as inducing cross-clade CTL responses. As early as 1988, CTLs specific for HIV-1 RT were found in blood samples from HIV-1 infected individuals (Hosmalin, A et al., 1990, Proc. Natl. Acad. Sci. USA, 87:2344-2348; Walker, BD et al., 1988, Science, 240:64-66.). Relatively strong Gag-specific CTL responses have been shown in numerous non-human primate and human studies, using DNA vaccines or a live recombinant vector containing viral Gag-Pol constructs (Evans, TG, 1999, J Infect Dis, 180:290-298; Ferrari, G, 1997, Blood, 90:2406-2416; Gorse, GJ et al., 1999, Vaccine, 18:835-849; Seth, A, 1998, Proc. Natl. Acad. Sci. USA, 95:10112-10116; Seth, A et al., 2000, J Virol, 74:2502-2509), but fewer Pol-specific CTL responses have been reported. The detection of significant CTL responses specific to Pol in our disclosure may be attributed in part to establishment of stable Pol expressing cell lines, in which codon alteration and inactivation of FS and PR in the Pol gene allow high level expression of the Pol protein without cellular toxicity. Though it remains possible that the hGag-Pol, or a combination of hGag and hPol, may exert similar effects with appropriate adjuvants or with different prime-boost regimens, the Rev-independent Gag-Pol fusion protein stimulates HIV-1 Gag and Pol specific CTL responses in mice and is envisioned as proving useful in an AIDS vaccine.

20

### PLASMID DESCRIPTIONS

#### Env Plasmids

##### VRC2100

25 pVR1012x/s R5gp139-Nef (delta) MHC (delta) CD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope protein gene from pX4gp160/h (Nabel lab #1272) was ligated in frame with the mutant Nef gene from pNefDMHCD4/h (Nabel lab #1278) to produce

30

pX4gp139-NefDMHCDCD4/h. The envelope-Nef fusion protein expressed from pX4gp139-NefDMHCDCD4/h contains the first 668 amino acids from the HIV envelope glycoprotein (gp139) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp139) contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain and regions important for oligomer formation. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. X4gp139-NefDMHCDCD4/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone.

#### 15 VRC2200

pVR1012x/s R5gp157-Nef (delta) MHC (delta) CD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (Nabel lab number 1272) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The envelope-Nef fusion protein expressed from pR5gp157-Nef/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (Nabel lab #1278) to produce pR5gp157-NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations

were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. R5gp157-NefDMHCDCD4/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone.

5

**VRC2300**

pVR1012x/s X4 gp139-Nef deltaMHC deltaCD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope protein gene from pX4gp160/h (Nabel lab #1272) was ligated in frame with the mutant Nef gene from pNefDMHC/h (Nabel lab #1276) to produce pX4gp139-NefDMHC/h. The envelope-Nef fusion protein expressed from pX4gp139-NefDMHC/h contains the first 668 amino acids from the HIV envelope glycoprotein (gp139) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp139) contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain and regions important for oligomer formation. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit MHC class I expression, point mutations were introduced into the Nef gene from pNef/h (Nabel lab #1275). The resulting amino acids substitutions in pNefDMHC/h are: P69A, P72A, P75A, and P78A. X4gp139-NefDMHC/h is expressed from the pVR1012x/s (Nabel lab #1267) vector backbone

30

**VRC2302**

pVR1012x/s X4gp130-Nef/h

35

For the X4gp130/h (VRC2703) portion, the protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons



optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 602. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Regions important for oligomer formation may be partially functional. This X4gp130/h (VRC2703) gene was fused to Nef gene. The Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create the viral Nef gene (pVR1012-Nef)(Nabel Lab #1093). The nucleotide sequence is homologous to the viral gene, and the protein encoded is the same. The expression vector backbone is pVR1012x/s (VRC2000).

15

**VRC2400**

pVR1012x/s X4gp157-NefDMHCDCD4/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The envelope-Nef fusion protein expressed from pX4gp157-NefDMHCDCD4/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The truncated envelope polyprotein (gp157) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (VRC3600) to produce pX4gp157-NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both

35

MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. X4gp160-NefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

#### VRC2700

pVR1012x/s X4gp140/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC2701

pVR1012x/s X4gp140(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking

the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC2702

pVR1012x/s X4gp128(del F/CL)/h

10

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 592. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

25

#### VRC2706

pVR1012x/s X4gp145/h

30

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704. The truncated envelope polyprotein contains the entire

35

SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

**VRC2707**

pVR1012x/s X4gp145(del F/CL del H IS)/h

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

15

20

25

**VRC2800**

pVR1012x/s R5gp140/h

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were

35

replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein  
5 contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).



**VRC2801**

pVR1012x/s R5gp140(del F/CL del H IS)/h

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC2804**

25 pVR1012x/s R5gp145/h

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

### VRC2805

10 pVR1012x/s R5gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

30

**VRC 2810**

pVR1012 x/s R5gp140delC1(delCFI)/h

5 Constant domain 1 was deleted from gp140delCFI from amino acid 33-127 and was replaced with a NheI site.

**VRC 2811**

pVR1012 x/s R5gp140delC2 (delCFI)/h

10 Constant domain 2 was deleted from gp140delCFI from amino acid 199-293, and was replaced with an NheI site

**VRC2812**

pVR 1012 x/s R5gp140delC3 (delCFI)/h

15

Constant domain 3 was deleted from gp140delCFI from amino acid 333-380, and was replaced with an NheI site.

**VRC2813**

20 pVR1012 x/s R5gp140delC4 (delCFI)/h

Constant domain 4 was deleted from gp140delCFI from amino acids 419-458, and was replaced with an NheI site.

25 **VRC2814**

pVR1012 x/s R5gp140delC5 (delCFI)/h

Constant domain 5 was deleted from gp140delCFI from amino acid 472 -498 and was replaced with an NheI site.

30

**VRC 2820**

pVR1012x/s R5gp140(dCFI)/dV1

35 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1 loop (a.a.129 to 154) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 2821**

20 pVR1012x/s R5gp140(dCFI)/dV2

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2 loop (a.a.160 to 193) and transmembrane domain.

Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2822**

pVR1012x/s R5gp140(dCFI)/dV3

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3 loop (a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2823**

pVR1012x/s R5gp140(dCFI)/dV4

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The



nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V4 loop (a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### **VRC 2824**

pVR1012x/s R5gp140(dCFI)/dV12

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2 loops (a.a.129 to 154, and a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required

for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2825

pVR1012x/s R5gp140(dCFI)/dV13

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3 loops (a.a.129 to 154, and a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2826

pVR1012x/s R5gp140(dCFI)/dV14

35 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V4 loops (a.a.129 to 154, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

## 20 VRC 2827

pVR1012x/s R5gp140(dCFI)/dV23

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion

domain, but lacking the V2, V3 loops (a.a.160 to 193, and a.a.299 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2828

10 pVR1012x/s R5gp140(dCFI)/dV24

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V4 loops (a.a.160 to 193, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2829**

pVR1012x/s R5gp140(dCFI)/dV34

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3, V4 loops (a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2830**

pVR1012x/s R5gp140(dCFI)/dV123

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were



replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.299 to 327) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2831

pVR1012x/s R5gp140(dCFI)/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V4 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for

oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2832

5 pVR1012x/s R5gp140(dCFI)/dV134

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V3,V4 loops (a.a. 129 to 154, a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2833

30 pVR1012x/s R5gp140(dCFI)/dV234

35 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L,

N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2,V3,V4 loops (a.a. 160 to 193, a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2834

pVR1012x/s R5gp140(dCFI)/dV1234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3,V4 loops (a.a. 129 to 154, a.a. 160 to 193, a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage

(F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

5

**VRC 2835**

pAdApt R5gp140(dCFI)/dV1

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1 loop (a.a.129 to 154) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

25

**VRC 2836**

30 pAdApt R5gp140(dCFI)/dV2

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the

35

protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2 loop (a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is AdApt.

#### VRC 2837

pAdApt R5gp140(dCFI)/dV3

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3 loop (a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been



deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

5     **VRC 2838**

pAdApt R5gp140(dCFI)/dV4

10     The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1  
15     envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein  
20     contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V4 loop (a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620,  
25     have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

**VRC 2839**

pAdApt R5gp140(dCFI)/dV12

30

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the  
35     protein encoded is the same with the following amino acid substitutions: F53L,

N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V2 loops (a.a.129 to 154, and a.a.160 to 193) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2840

pAdApt R5gp140(dCFI)/dV13

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V3 loops (a.a.129 to 154, and a.a.299 to 327) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2,

from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2841

5 pAdApt R5gp140(dCFI)/dV14

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1, V4 loops (a.a.129 to 154, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2842

pAdApt R5gp140(dCFI)/dV23

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an

35

R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length

5 R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V3 loops (a.a.160 to 193, and a.a.299 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required

10 for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

15 **VRC 2843**

pAdApt R5gp140(dCFI)/dV24

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

20 gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1

25 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein

30 contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2, V4 loops (a.a.160 to 193, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from

amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2844

5 pAdApt R5gp140(dCFI)/dV34

10 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V3, V4 loops (a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2845

pAdApt R5gp140(dCFI)/dV123

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an

35



R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length

5 R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.299 to 327) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS)

10 are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt

15

VRC 2846

pAdApt R5gp140(dCFI)/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

20 gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an

25 R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length

30 R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V4 loops (a.a. 129 to 154, a.a.160 to 193,and a.a.386 to 413) and transmembrane domain. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains,

35 from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad

(H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

5      VRC 2847

pAdApt R5gp140(dCFI)/dV134

10      The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1  
15      envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein  
20      contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V3,V4 loops (a.a. 129 to 154, a.a.299 to 327, and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1  
25      and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

30      VRC 2848

pAdApt R5gp140(dCFI)/dV234

35      The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the

protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V2,V3,V4 loops (a.a. 160 to 193, a.a.299 to 327,and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

#### VRC 2849

pAdApt R5gp140(dCFI)/dV1234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 680. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the V1,V2,V3,V4 loops (a.a. 129 to 154, a.a. 160 to 193, a.a.299 to 327,and a.a.386 to 413) and transmembrane domain. Heptad(H) 1, Heptad

2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pAdApt.

**VRC2850**

pVR1012 x/s R5gp145delC1 (delCFI)/h

Constant domain 1 was deleted from gp145delCFI from amino acid 33-127 and was replaced with an Nhe I site.

**VRC2851**

pVR1012 x/s R5gp145delC2 (delCFI)/h

Constant domain 2 was deleted from gp145dCFI from amino acid 199-293 and was replaced with an Nhe I site.

**VRC2852**

pVR1012 x/s R5gp145delC3 (delCFI)/h

Constant domain 3 was deleted from gp145delCFI from amino acid 333-380 and was replaced with an Nhe I site.

**VRC2853**

pVR1012 x/s R5gp145delC4 (delCFI)/h

Constant domain 4 was deleted from gp145delCFI from amino acid 419-458 and was replaced with an Nhe I site.

**VRC2854**

pVR1012 x/s R5gp145delC5 (delCFI)/h

Constant domain 5 was deleted from gp145delCFI from amino acid 472-498 and was replaced with an Nhe I site.

**VRC 2860**

pVR1012x/s R5gp145(dCFI)/h/dV1

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V1 loop (a.a. 129-154) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2861**

pVR1012x/s R5gp145(dCFI)/h/dV2

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with



the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V2 loop (a.a. 160-193) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2862

pVR1012x/s R5gp145(dCFD)/h/dV3

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V3 loop (a.a. 299-327) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2863**

pVR1012x/s R5gp145(dCFI)/h/dV4

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V4 loop (a.a. 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2864**

25 pVR1012x/s R5gp145(dCFI)/h/dV12

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

35

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V1, V2 loops (a.a. 129-154 and 160-193) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2865

pVR1012x/s R5gp145(dCFI)/h/dV13

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V3 loops (a.a. 129-154 and 299-327) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

VRC 2866

pVR1012x/s R5gp145(dCFI)/h/dV14

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V4 loops (a.a. 129-154 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

VRC 2867

25 pVR1012x/s R5gp145(dCFI)/h/dV23

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

35

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V2, V3 loops (a.a. 160-193 and 299-327) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2868

pVR1012x/s R5gp145(dCFI)/h/dV24

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V2, V4 loops (a.a. 160-193 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).



**VRC 2869**

pVR1012x/s R5gp145(dCFI)/h/dV34

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V3, V4 loops (a.a. 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 2870**

25 pVR1012x/s R5gp145(dCFI)/h/dV134

30 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession

35

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V3, V4 loops (a.a. 129-154, 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2871

pVR1012x/s R5gp145(dCFI)/h/dV234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The V2, V3, V4 loops (a.a. 160-193, 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

VRC 2872

pVR1012x/s R5gp145(dCFI)dv123/h

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version  
10 of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated  
15 after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V4 loops (a.a. 129-154, 160-193 and 386-413) and Fusion and  
20 Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

VRC 2873

25 pVR1012x/s R5gp145(dCFI)/h/dV124

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The  
30 nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with  
35 the corresponding region from the BaL strain of HIV-1 (GeneBank accession

number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V4 loops (a.a. 129-154, 160-193 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 2874

pVR1012x/s R5gp145(dCFI)/h/dV1234

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, and P470L. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was terminated after the codon for amino acid 704. The truncated envelope polyprotein (gp145) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad (H) 1, Heptad 2 and their Interspace (IS) are required for oligomerization. The V1, V2, V3, V4 loops (a.a. 129-154, 160-193, 299-327 and 386-413) and Fusion and Cleavage (F/CL) domains (a.a. 503-536) have been deleted. Also, the Interspace (IS) between Heptad (H) 1 and 2 (a.a.593-620) have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC2900**

pVR1012x/s R5gp150/h

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an  
10 R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The full-length R5-tropic version of the envelope protein gene from pR5gp160/h (VRC3000) was  
15 terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC3000**

pVR1012x/s R5gp160/h

20 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The  
25 nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1  
30 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.



**VRC3200**

pVR1012x/s X4gp150/h

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. The expression vector backbone is pVR1012x/s (VRC2000).

15

**VRC3201**

pVR1012x/s X4gp150(del F/CL del H IS)/h

20 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. The full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. The truncated envelope polyprotein (gp150) contains the entire SU protein and a portion of the TM protein including the fusion domain, the transmembrane domain, and regions important for oligomer formation. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

30

**VRC3202**

pVR1012x/s X4gp150 Δgly/h.

5 Eukaryotic vector with humanized codons expressing the HIV envelope glycoprotein gp150 from HXB2, X4 tropic mutated in the Glycosylation sites. VRC3202 pVR1012x/s X4gp150Dgly/h The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4- tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows

10 little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycosylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting

15 amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, N295D, N332D, N339D, N356D, N386D, and N448D. The full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for

20 oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

**VRC3203**

pVR1012x/s X4gp150 AB Δgly/h

25

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The AB designation means the amino acids from 1-307 (XbaI to EcoRI). The nucleotide

30 sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycosylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-length X4-tropic version of the envelope protein gene from pX4gp160/h

(VRC3300). The resulting amino acids substitutions in X4gp160Dgly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, and N295D. The full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 752.

5 Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

#### VRC3300

10 pVR1012x/s X4gp160/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The

15 nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

20

#### VRC3301

pVR1012x/s X4gp160(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the

25 gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. Full length SU

30 and TM-proteins are Expressed. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

35

**VRC3400**pVR1012x/s X4gp160 $\Delta$ gly/h

5 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycosylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting amino acids substitutions in X4gp160 $\Delta$ gly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, N295D, N332D, N339D, N356D, N386D, and N448D. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

**VRC3401**pVR1012x/s X4gp160AB mut  $\Delta$ gly/h

20 The protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The AB designation means the amino acids from 1-307 (XbaI to EcoR1). The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To disrupt potential glycosylation sites in the HIV-1 envelope proteins, point mutations were introduced into the full-length X4-tropic version of the envelope protein gene from pX4gp160/h (VRC3300). The resulting amino acids substitutions in X4gp160 $\Delta$ gly/h are: N88D, N156D, N160D, N197E, N230D, N234D, N241D, N276D, L288V, N289D, S291T, and N295D. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. Full length SU and TM proteins are expressed from the pVR1012x/s (VRC2000) vector backbone.

**VRC3500**

pVR1012x/s Nef/h

5 The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. The nucleotide sequence Nef/h shows little homology to the viral gene, but the protein encoded is the same. Nef/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

10

**VRC3600**

pVR1012x/s NefDMHCDCD4/h

15

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. pNefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

20

**VRC3700**

pVR1012x/s NefDCD4/h

25

The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDCD4/h are: D174A and D175A. pNefDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone. VRC3700 pVR1012x/s NefDCD4/h The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit CD4 expression,

30

35



point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDCD/h are: D174A and D175A. pNefDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

5      **VRC3800**

pVR1012x/s NefDMHC/h

10      The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit MHC class I expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHC/h are: P69A, P72A, P75A, and P78A. pNefDMHC/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

15

**VRC5200**

pVR1012x/s 89.6Pgp128(del F/CL)/h

20      The protein sequence of the envelope polyprotein (gp160) from 89.6P (Dual-tropic, GenBank accession number u89134/LOCUS:SIU89134) was used to create a synthetic version of the gene (89.6Pgp160/h) using codons optimized for expression in human cells. The nucleotide sequence 89.6Pgp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-length 89.6P, dual-tropic version of the envelope protein gene from 89.6P gp160/h (VRC3000) was  
25      terminated after the codon for amino acid 596. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

30

**VRC5201**

pVR1012x/s 89.6Pgp140(del F/CL del H IS)/h

35      The protein sequence of the envelope polyprotein (gp160) from 89.6P (dual-tropic, GenBank accession number u89134/locus SIU89134) was used to create a synthetic

version of the gene (Dualtropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence dualtropic gp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-length 89.6P, dual-tropic version of the envelope protein gene was terminated after the codon for amino acid 683. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 597-625, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5202

pVR1012x/s 89.6Pgp145(del F/CL del H IS/h

The protein sequence of the envelope polyprotein (gp160) from 89.6P (dual-tropic, GenBank accession number u89134/locus SIU89134) was used to create a synthetic version of the gene (Dualtropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence dualtropic gp160/h shows little homology to the 89.6P gene, but the protein encoded is the same. The full-length 89.6P, dual-tropic version of the envelope protein gene was terminated after the codon for amino acid 709. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 508-541, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 597-625, have been deleted. Regions important for oligomer formation may be partially functional. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5203**

pVR1012x/s 89.6Pgp160/h

5 The protein sequence of the envelope polyprotein (gp160) from 89.6P (Dual-tropic, GenBank accession number U89134/LOCUS: SIU89134) was used to create a synthetic version of the gene (dual tropic gp160/h) using codons optimized for expression in human cells. The nucleotide sequence 89.6P gp160/h shows little  
10

**VRC5300**

pVR1012x/s R5(clade C)gp140(del F/CL del H IS)/h

15 The protein sequence of the envelope polyprotein (gp160) from 92br025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (R5gp160/h) using codons optimized for expression in human cells. The nucleotide sequence R5gp160/h shows little homology to the 92br025 gene, but the protein encoded is the same. The full-length R5-tropic version of the envelope protein was synthesized by Operon under the name: kongene. The XbaI (18nt up-  
20 stream from ATG) to BglII (1376nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BglII sites of VRC2701 pVR1012x/s X4gp140(del F/CL del H IS)/h backbone. Therefore, the gene is R5 (clade C) gp160/h up to the BglII site (1376nt from ATG) and the rest of the gene after BglII site is VRC2701 pVR1012x/s  
25 X4gp140(del F/CL del H IS)/h. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids  
30 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000). The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 680.

**VRC5301**

pVR1012x/s R5(clade C)gp145(del F/CL del H IS)/h

5 The protein sequence of the envelope polyprotein (gp160) from 92br025 (R5-tropic, GenBank accessionnumber U52953) was used to create a synthetic version of the gene (R5gp160/h) using codons optimized for expression in human cells. The nucleotide sequence R5gp160/h shows little homology to the 92br025 gene, but the protein encoded is the same. The full-length R5-tropic version of the envelope protein was synthesized by Operon under the name: kongene. The XbaI (18nt up-  
10 stream from ATG) to BglII (1376nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BglII sites of VRC2701 pVR1012x/s X4gp140(del F/CL del H IS)/h backbone. Therefore, the gene is R5 (clade C) gp160/h up to the BglII site (1376nt from ATG) and the rest of the gene after BglII site is VRC2707 pVR1012x/s  
15 X4gp145(del F/CL del H IS)/h. The truncated envelope polyprotein contains the entire SU protein and a portion of the TM protein including the fusion domain, but lacking the transmembrane domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids  
20 503-536, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 593-620, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000). The full-length X4-tropic version of the envelope protein from pX4gp160/h (VRC3300) was terminated after the codon for amino acid 704.

25

**VRC5303**

pVR1012x/s R5gp145CladeC(Brazil) delCFI/h

Authentic Clade C unlike 5301 which is a hybrid between C and B.

30

**VRC 5304**

pVR1012x/s R5(clade A)gp140(del F/CL del H IS)/h

35 The protein sequence of the envelope polyprotein (gp160) from 92rw020 (R5-tropic, GenBank accessionnumber U51283) was used to create a synthetic version of the

gene (Clade-A gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the 92rw020 gene, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1837 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 486-519, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 576-604, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5305

pVR1012x/s R5(clade A)gp145(del F/CL del H IS)/h

The protein sequence of the envelope polyprotein (gp160) from 92rw020 (R5-tropic, GenBank accessionnumber U51283) was used to create a synthetic version of the gene (Clade-A gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the 92rw020 gene, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1912 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 486-519, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 576-604, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).



**VRC 5306**

pVR1012x/s R5(clade E)gp140(del F/CL del H IS)/h

5 The protein sequence of the envelope polyprotein (gp140delCFI) from 93th966.8 (R5-tropic, GenBank accessionnumber U08456) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 93th966.8, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1856 nt down-stream from ATG) fragment which contains  
10 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for  
15 oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 497-530, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 588-613, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

20 **VRC 5307**

pVR1012x/s R5(clade E)gp145(del F/CL del H IS)/h

25 The protein sequence of the envelope polyprotein (gp145delCFI) from 93th966.8 (R5-tropic, GenBank accessionnumber U08456) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 93th966.8, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1928 nt down-stream from ATG) fragment which contains  
30 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for  
35 oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 497-530, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from

amino acids 588-613, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5308

5 pVR1012x/s R5(clade C South African)gp140(del F/CL del H IS)/h

10 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accessionnumber AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1833 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein  
15 contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 487-520, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from  
20 amino acids 577-605, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5309

25 pVR1012x/s R5(clade C South African)gp145(del F/CL del H IS)/h

30 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accessionnumber AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein  
35 contains the entire SU protein and the TM domain, but lacking the the Fusion domain and Cytoplasmic domain. Regions important for oligomer formation may be

partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The Fusion and Cleavage (F/CL) domains, from amino acids 487-520, have been deleted. The Interspace (IS) between Heptad (H) 1 and 2, from amino acids 577-605, have been deleted. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5350

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1 loop(a.a.133-148) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5351

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV12

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into

the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2 loops (a.a.133-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

10 **VRC 5352**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV123

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V3 loops (a.a.130-191, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

30 **VRC 5353**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV1234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V3,V4 loops (a.a.130-191, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5354

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV124

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2,V4 loops (a.a. 130-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).



VRC 5355

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV13

5 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little  
10 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3 loops (a.a.133-148, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the  
15 Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

20 VRC 5356

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV134

25 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into  
30 the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3,V4 loops (a.a.130-148, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions  
35 important for oligomer formation may be partially functional. Heptad(H) 1, Heptad

2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5357

5 pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV14

10 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little  
15 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope  
20 polyprotein contains deletion in the V1,V4 loops (a.a. 130-148, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their  
25 Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5358

25 pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV2

30 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little  
35 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2 loop(a.a.154-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS)

between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5359

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV23

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V3 loops (a.a. 154-191, 330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5360

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV234

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into

the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V3,V4 loops (a.a.154-191, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

10 **VRC 5361**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV24

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V4 loops (a.a. 154-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

30 **VRC 5362**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV3

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3 loop(a.a.330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

### VRC 5363

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV34

15

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little  
20  
homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3,V4 loops (a.a.330-358, 384-408) of the SU  
25  
protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is  
30  
pVR1012x/s (VRC2000).



**VRC 5364**

pVRC1012(x/s)-gp140(dCFI)(Brazil C)/dV4

5 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp140(delCFI) shows little  
10 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V4 loop(a.a.384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), transmembrane domain, and  
15 cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 5365**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1

25 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into  
30 the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1 loop(a.a.133-148) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional.

Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5366

5 PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV12

10 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
15 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope  
20 polyprotein contains deletion in the V1,V2 loops (a.a.133-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The  
expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5367

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV123

25 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
30 homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope  
polyprotein contains deletion in the V1, V2,V3 loops (a.a. 133-191,330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529),  
35 the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic

domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5     **VRC 5368**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV1234

10     The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
15     homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V2, V3,V4 loops (a.a. 133-191,330-358,384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and  
20     Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 5369**

25     PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV124

30     The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
35     homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V2,V4 loops (a.a. 133-148,154-191,384-

408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5370

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV13

10

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
15  
homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1,V3 loops (a.a.133-148, 330-358) of the SU  
20  
protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

25

#### VRC 5371

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV134

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little  
30  
homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into  
35

the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V3, V4 loops (a.a. 133-148, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

10 **VRC 5372**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV14

15 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(dCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into 20 the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V1, V4 loops (a.a. 133-148, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially 25 functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 5373**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV2

30

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(dCFI) shows little 35 homology to the gene 92BR025, but the protein encoded is the same. The XbaI



(18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2 loop (a.a.154-191) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

10

**VRC 5374**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV23

15

20

25

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2,V3 loops (a.a.154-191,330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

30

**VRC 5375**

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV234

35

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little

homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2, V3, V4 loops (a.a. 154-191, 330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5376

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV24

15

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V2, V4 loops (a.a. 154-191, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains (a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2 (a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

25

#### VRC 5377

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV3

30

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version

of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3 loop (a.a.330-358) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC 5378

PVRC1012(x/s)-gp145(dCFI)(Brazil C)/dV34

The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1910 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V3,V4 loops (a.a.330-358, 384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 5379**

PVR1012(x/s)-gp145(dCFI)(Brazil C)/dV4

5 The protein sequence of the envelope polyprotein (gp160) from 92BR025 (R5-tropic, GenBank accession number U52953) was used to create a synthetic version of the gene (Brazil-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence of Brazil-C gp145(delCFI) shows little homology to the gene 92BR025, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1910 nt down-stream from ATG) fragment  
10 which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein contains deletion in the V4 loop (a.a.384-408) of the SU protein. It also lacks the Fusion and Cleavage (F/CL) domains( a.a. 496-529), the Interspace (IS) between Heptad (H) 1 and 2(a.a. 586-612), and Cytoplasmic domain in the TM  
15 protein. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5500**

20 pVR1012x/s R5(SA-C )gp140(dCFI) dV1/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression  
25 in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
30 the V1 loop (a.a.136-150), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

35

**VRC5501**

pVR1012x/s R5(SA-C )gp140(dCFI) dV12/h

5 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains  
10 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2 loops (a.a.136-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer  
15 formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC 5502**

20 pVR1012x/s R5(SA-C )gp140(dCFI) dV123/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression  
25 in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
30 the V1, V2, V3 loops (a.a. 136-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is  
35 pVR1012x/s (VRC2000).



**VRC5503**

pVR1012x/s R5(SA-C)gp140(dCFI)dV1234/h

5 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream  
10 from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V3, V4 loops (a.a.136-194,297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2  
15 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5504**

pVR1012x/s R5(SA-C)gp140(dCFI)dV124/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic  
25 version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
30 BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V4 loops (a.a. 136-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their

Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5505

5 pVR1012x/s R5(SA-C)gp140(dCFI)dV13/h

10 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks 15 the V1, V3 loops (a.a.136-150, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is 20 pVR1012x/s (VRC2000).

#### VRC5506

pVR1012x/s R5(SA-C)gp140(dCFI)dV134/h

25 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks 30 the V1, V3, V4 loops (a.a. 136-150, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions 35

important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5     **VRC5507**

pVR1012x/s R5(SA-C)gp140(dCFI)dV14/h

10     The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V4 loops (a.a.136-150, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

15     **VRC5508**

pVR1012x/s R5(SA-C)gp140(dCFI)dV2/h

25     The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2 loop (a.a.156-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane

30

35

domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5      **VRC5509**

pVR1012x/s R5(SA-C)gp140(dCFI)dV23/h

10      The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
15      BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3 loops (a.a.156-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their  
20      Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5510**

pVR1012x/s R5(SA-C)gp140(dCFI)dV234/h

25

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to  
30      the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3, V4 loops (a.a. 156-194, 297-325, 385-399), the Fusion and Cleavage  
35      (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2

(a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

**VRC5511**

pVR1012x/s R5(SA-C)gp140(dCFI)dV24/h

10 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains  
15 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V4 loops (a.a.156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer  
20 formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5512**

25 pVR1012x/s R5(SA-C)gp140(dCFI)dV3/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression  
30 in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
35 the V3 loop (a.a.297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520),



the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

**VRC5513**

pVR1012x/s R5(SA-C)gp140(dCFI)dV34/h

10 The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains  
15 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3, V4 loops (a.a. 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer  
20 formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5514**

25 pVR1012x/s R5(SA-C)gp140(dCFI)dV4/h

The protein sequence of the envelope polyprotein (gp140delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp140delCFI) using codons optimized for expression  
30 in human cells. The nucleotide sequence R5gp140delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
35 the V4 loop (a.a 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520),

the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), the transmembrane domain and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5

**VRC5515**

pVR1012x/s R5(SA-C)gp145(dCFI)dV1/h

10 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains  
15 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1 loop (a.a.136-150), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional.  
20 Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5516**

pVR1012x/s R5(SA-C)gp145(dCFI)dV12/h

25

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to  
30 the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1,V2 loops (a.a.136-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-  
35 520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the

intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5     **VRC5517**

pVR1012x/s R5(SA-C)gp145(dCFI)dV123/h

10     The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
15     BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V2, V3 loops (a.a.136-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for  
20     oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**VRC5518**

pVR1012x/s R5(SA-C)gp145(dCFI)dV1234/h

25     The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream  
30     from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1,V2,V3 ,V4 loops (a.a.136-194, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2  
35     (a.a.577-605), and the intracellular region. Regions important for oligomer

formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5        **VRC5519**

pVR1012x/s R5(SA-C)gp145(dCFI)dV2/h

10        The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
15        BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2 loop (a.a.156-194), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The  
20        expression vector backbone is pVR1012x/s (VRC2000).

**VRC5520**

pVR1012x/s R5(SA-C)gp145(dCFI)dV23/h

25        The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
30        BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3 loops (a.a.156-194, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and  
35        the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5521

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV234/h

The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V3, V4 loops (a.a.156-194, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5522

pVR1012x/s R5(SA-C)gp145(dCFI)dV24/h

25 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V2, V4 loops (a.a.156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially



functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5523

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV3/h

10 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks 15 the V3 loop (a.a.297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

20

#### VRC5524

pVR1012x/s R5(SA-C)gp145(dCFI)dV34/h

25 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains 30 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V3, V4 loops (a.a. 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5525

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV4/h

10 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
15 the V4 loop (a.a.385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

20

#### VRC5526

pVR1012x/s R5(SA-C)gp145(dCFI)dV13/h

25 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains  
30 polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V3 loops (a.a.136-150, 297-325), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially

functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

#### VRC5527

5 pVR1012x/s R5(SA-C)gp145(dCFI)dV134/h

10 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
15 the V1, V3, V4 loops (a.a.136-150, 297-325, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is  
20 pVR1012x/s (VRC2000).

#### VRC5528

pVR1012x/s R5(SA-C)gp145(dCFI)dV124/h

25 The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamHI (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to BamHI sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks  
30 the V1, V2, V4 loops (a.a.136-150, 156-194, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer  
35

formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

5      **VRC5529**

pVR1012x/s R5(SA-C)gp145(dCFI)dV14/h

10      The protein sequence of the envelope polyprotein (gp145delCFI) from 97ZA012 (R5-tropic, GenBank accession number AF286227) was used to create a synthetic version of the gene (Clade-C gp145delCFI) using codons optimized for expression in human cells. The nucleotide sequence R5gp145delCFI shows little homology to the gene 97ZA012, but the protein encoded is the same. The XbaI (18nt up-stream from ATG) to BamH1 (1914 nt down-stream from ATG) fragment which contains polylinker at the 5' end, Kozak sequence and ATG was cloned into the XbaI to  
15      BamH1 sites of pVR1012x/s backbone. The truncated envelope polyprotein lacks the V1, V4 loops (a.a.136-150, 385-399), the Fusion and Cleavage (F/CL) domains (a.a. 487-520), the Interspace (IS) between Heptad (H) 1 and 2 (a.a.577-605), and the intracellular region. Regions important for oligomer formation may be partially functional. Heptad(H) 1, Heptad 2 and their Interspace(IS) are required for  
20      oligomerization. The expression vector backbone is pVR1012x/s (VRC2000).

**Gag-Pol Plasmids**

**VRC3900**

pVR1012x/s Gag/h

25

30      The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein to produce pGag(fs)Pol/h (VRC4200). The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To produce a gene that

expresses all of the gag proteins, the region encoding pol amino acids 77-1003 were deleted from Gag(fs)Pol/h to produce Gag/h. This construct also encodes most of the protease (98 amino acids) gene encoding amino acids 3-77. Gag/h is expressed from the pVR1012x/s vector backbone.

5

**VRC3901**

pVR1012x/s SIV Gag/h

10 The protein sequence of the gag polyprotein (amino acid from 1-550) SIVmac239 (GenBank accession number M33262) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the SIVMac239 gene, but the protein encoded is the same.

15

**VRC4000**

pVR1012x/s Gag-Pol/h

20 The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene  
25 was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. Gag-Pol/h is expressed from the pVR1012x/s vector backbone.

30



**VRC4001**

pVR1012x/s SIVGag-Pol/h

5 Eukaryotic vector with humanized codons expressing the Gag-pol gene of SIVmac239. The Sal I –XbaI fragment of SIV gag(VRC3901) was inserted into Sal I –XbaI of SIV Pol(VRC4101) to create VRC4001.

**VRC4100**

pVR1012x/s Pol/h

10

The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To initiate translation at the beginning of Pol, a methionine codon was added to the 5'-end of the synthetic polymerase gene to create the Pol/h gene. Pol/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

15

**VRC4101**

pVR1012x/s SIV Pol/h

20

The protein sequence of the pol polyprotein (amino acids 3-1017) from SIVmac239 (GenBank accession number M19921) was used to create a synthetic version of the pol gene using codons optimized for expression in human cells. To initiate translation at the beginning of Pol, a methionine codon was added to the 5'-end of the synthetic polymerase gene to create the Pol/h gene. The Protease (Pr) mutation is at pol amino acid 123 and is AAG->GGA or amino acids R->G. Reverse transcriptase(RT) mutation is at aa 352(GAC to CAT D(D to H) –RT) and Integrase mutation is at aa 788(GAC to GGC(D to A) –IN). Pol/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

25

30

**VRC4200**

pVR1012x/s Gag(fs)Pol/h

35 The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the

gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone.

#### VRC4300

20 pVR1012 Gag-Pol(d delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame shifted Pol genes of HIV HXB2 subtype B with deletions in Reverse transcriptase, and Integrase regions. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express

the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

### VRC4301

pVR1012x/s-Gag(FS)-Pol-delta RT IN-IRES-R5 gp157-Nef

For the Gag(FS)Pol delta RT delta IN/h portion, VRC4302 pVR1012x/s Gag(fs)Pol(delta PR delta/h was used for the protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy- terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid

771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. This gene has been fused to an Internal Ribosomal Entry Site (IRES) and then fused to the R5gp157-Nef from VRC2200 pVR1012x/s R5gp157- NefDMHCDCD4/h in which the protein sequence of the envelope polyprotein (gp160) from HXB2 (X4-tropic, GenBank accession number K03455) was used to create a synthetic version of the gene (X4gp160/h) using codons optimized for expression in human cells. The nucleotide sequence X4gp160/h shows little homology to the HXB2 gene, but the protein encoded is the same with the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T. To produce an R5-tropic version of the envelope protein (R5gp160/h), the region encoding HIV-1 envelope polyprotein amino acids 275 to 361 from X4gp160/h (VRC3300) were replaced with the corresponding region from the BaL strain of HIV-1 (GeneBank accession number M68893, again using human preferred codons). The envelope-Nef fusion protein expressed from pR5gp157-Nef/h contains the first 820 amino acids from the HIV envelope glycoprotein (gp157) fused to the entire mutant Nef protein. The gene for gp157 was ligated in frame with the full-length mutant Nef gene from pNefDMHCDCD4/h (VRC3600) to produce pR5gp157- NefDMHCDCD4/h. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. To disrupt the ability of Nef to limit both MHC class I and CD4 expression, point mutations were introduced into the Nef gene from pNef/h (VRC3500). The resulting amino acids substitutions in pNefDMHCDCD4/h are: P69A, P72A, P75A, P78A, D174A and D175A. R5gp157-NefDMHCDCD4/h is expressed from the pVR1012x/s (VRC2000) vector backbone.

#### VRC4302

pVR1012 Gag(delFS)Pol(delta PR delta RT delta IN)/h

30

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag

35

proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A.

Note: This vector is not in the pVR1012x/s backbone.

#### VRC4303

pVR1012 SIV Gag(delFS)Pol(delta PR delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame-shift-deleted Pol genes SIVmac239 with deletions in the Protease, Reverse transcriptase, and Integrase regions. The 5 of Ts from 3188 to 3192 of SIV gag-pol(VRC4001) were deleted to create VRC4303.

#### VRC4304

pVR1012 Gag (delFS)Pol delta PR delta RT delta IN/h

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 C clade (GenBank accession number U52953) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The



nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

#### VRC4305

25 pVR1012 Gag-A(delFS)Pol(delta PR delta RT delta IN)/h

Eukaryotic vector with humanized codons expressing the Gag and the frame shifted Pol genes of HIV HIV-1A clade with deletions in the Protease, Reverse transcriptase, and Integrase regions. VRC4305 pVR1012 Gag(-AdelFS)Pol(delta PR delta RT delta IN)/h The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 A clade (GenBank accession number AF004885) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally

cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

20

**pVRC4306**

pVR1012 Gag(delFS)Pol delta PR delta RT delta IN/Nef/h

The protein sequence of the Gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the Gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic Gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic Gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the Gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic Gag gene was ligated in frame with sequences encoding the Pol polyprotein. The protein sequence of the Pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the Pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the Gag-Pol

35

polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of Gag plus an additional 3 amino acids from Pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all Gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. The Protease (PR) mutation is at Gag-Pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at Gag-Pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at Gag- Pol amino acid 1209 and is ACT->CAT or amino acids D->A. The Nef/h gene was fused to downstream of Pol gene of Gag(delFS)Pol(delta PR delta RT delta IN). No loss or extra-amino acid was created by the fusion between Nef and Pol. The ATG of Nef was preserved. The protein sequence of the Nef protein from HIV-1 PV22 (GenBank accession number K02083) was used to create a synthetic version of the Nef gene (Nef/h) using codons optimized for expression in human cells. The nucleotide sequence Nef/h shows little homology to the viral gene, but the protein encoded is the same. Note: This vector is not in the pVR1012x/s backbone.

20

**VRC 4308**

VRC4302-myr

pVR1012 Gag (delFS) Pol deltaPR deltaRT deltaIN deltaMyr/h

25

The myristoylation site was deleted from pVRC4302. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons

30

35

optimized for expression in human cells. To create the possibility for translational  
 frameshifting as means to express the gag-pol polyprotein, the synthetic coding  
 region for the last four amino acids of the NC protein through the rest of gag plus an  
 additional 3 amino acids from pol were replaced with the corresponding viral  
 sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank  
 accession number M19921). The substitution of viral for synthetic sequences both  
 introduces the sites required for frameshifting and restores the ability to express all  
 gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T  
 nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed  
 from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol  
 amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse  
 Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or  
 amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and  
 is ACT->CAT or amino acids D->A. The myristylation site was deleted. Note: This  
 vector is not in the pVR1012x/s backbone.

#### VRC 4309

pVR1012 Gag (delFS) Pol deltaPR deltaRT deltaIN delta Myr/Nef/h

The myristoylation site was deleted from pVRC4306. The protein sequence of the  
 gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number  
 K03455) was used to create a synthetic version of the gag gene using codons  
 optimized for expression in human cells. The nucleotide sequence of the synthetic  
 gag gene shows little homology to the HXB2 gene, but the protein encoded is the  
 same. The synthetic gag gene contains all of the mature Gag proteins except for the  
 last two that are normally cleaved from the carboxy-terminus of the gag polyprotein,  
 p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with  
 sequences encoding the pol polyprotein. The protein sequence of the pol  
 polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number  
 M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons  
 optimized for expression in human cells. To create the possibility for translational  
 frameshifting as means to express the gag-pol polyprotein, the synthetic coding  
 region for the last four amino acids of the NC protein through the rest of gag plus an  
 additional 3 amino acids from pol were replaced with the corresponding viral  
 sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank

accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic B clade Nef ( Genbank access number ) was fusion to the 3' end of pol by PCR. Note: This vector is not in the pVR1012x/s backbone.

#### VRC 4310

pVR1012 Nef Gag (del fs) (del Myr) Pol (delta PR delta RT delta IN)/h

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector



backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A.

5 The stop codon TAG of synthetic B clade Nef ( Genbank access number ) gene was removed and was fused to the 5' end of gene gene by PCR. Note: This vector is not in the pVR1012x/s backbone.

### VRC 4311

10 pVR1012 Gag-C(delFS)Pol(delta PR delta RT delta IN) Nef/h  
VRC4304 + Clade C Nef

The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 C clade (GenBank accession number U52953) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500).

15 The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921).

20 The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN)

25

30

35

mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic C clade Nef ( Genbank accession number:U52953) was fusion to the 3' end of pol by PCR. Note:This vector is not in the pVR1012x/s backbone.

5

### VRC 4312

Gag (del fs) (del Myr) Nef Pol ΔPRΔRTΔIN/h

VRC-Myr-gag-Dnef-Dpol

(Eukaryotic vector with humanized codons expressing the Gag, truncated Nef, and truncated Pol proteins of HIV subtype B .ns.)

10

This construct was derived from VRC4302. The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HXB2 (GenBank accession number K03455) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HXB2 gene, but the protein encoded is the same. The synthetic gag gene contains all of the mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding synthetic nef gene that 51 aa were deleted from 5'. 77 aa were deleted from 5' of pol polyprotein, and ligated with 3' of nef in which tag stop codon was deleted. The protein sequence of the pol polyprotein (amino acids 3-78 - 1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG-

15

20

25

30

>GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. Note: This vector is not in the pVR1012x/s backbone.

5

**VRC 4313**pVR1012 Gag Clade A (del fs)Pol( $\Delta$  PR  $\Delta$  RT  $\Delta$  IN)/h

VRC4305+ Clade A Nef

- 10 The protein sequence of the gag polyprotein (Pr55, amino acids 1-432) from HIV-1 A clade (GenBank accession number AF004885) was used to create a synthetic version of the gag gene using codons optimized for expression in human cells. The nucleotide sequence of the synthetic gag gene shows little homology to the HIV-1 gene, but the protein encoded is the same. The synthetic gag gene contains all of the
- 15 mature Gag proteins except for the last two that are normally cleaved from the carboxy-terminus of the gag polyprotein, p1 and p6 (amino acids 433-500). The synthetic gag gene was ligated in frame with sequences encoding the pol polyprotein. The protein sequence of the pol polyprotein (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) was used to create a synthetic version
- 20 of the pol gene (Pol/h) using codons optimized for expression in human cells. To create the possibility for translational frameshifting as means to express the gag-pol polyprotein, the synthetic coding region for the last four amino acids of the NC protein through the rest of gag plus an additional 3 amino acids from pol were replaced with the corresponding viral sequences (nucleotides 2074-2302 on the
- 25 HXB2 genome) from NL4-3 (GenBank accession number M19921). The substitution of viral for synthetic sequences both introduces the sites required for frameshifting and restores the ability to express all gag proteins, including p1 and p6. The deleted Frame Shifted (delFS) has 5 T nucleotides deleted between the Gag and Pol sequences. Gag(fs)Pol/h is expressed from the pVR1012x/s vector
- 30 backbone. The Protease (PR) mutation is at gag-pol amino acid 553 and is AGG->GGC or amino acids R->G. The Reverse Transcriptase (RT) mutation is at gag-pol amino acid 771 and is GAC->CAC or amino acids D->H. The Integrase (IN) mutation is at gag-pol amino acid 1209 and is ACT->CAT or amino acids D->A. The stop codon TAG was removed and synthetic A clade Nef ( Genbank accession

number:AF069670) was fused to the 3' end of pol by PCR. Note: This vector is not in the pVR1012x/s backbone.

### Examples

5

#### *Env*

##### **Immunogens**

Plasmids expressing the CXCR4-tropic HIV-1 HXB2 Env were made synthetically with sequences designed to disrupt viral RNA structures that limit protein expression by using codons typically found in human cells. Briefly, the synthetic Env gene of HXB2 (GenBank accession number K03455) was generated in three fragments by assembling the overlapping synthetic oligonucleotides using PCR amplification. To produce a CCR5-tropic version of the HIV-1 envelope, the region encoding amino acids 275 to 361 of HXB2 (CXCR4-tropic) gp160 was replaced with CCR5-tropic HIV-1 BaL sequence (GenBank accession number M68893), which includes the V3 loop. Glycosylation mutants were generated by site-directed mutagenesis to replace asparagine with glutamic acid residues at seventeen conserved glycosylation sites between amino acids 88 and 448. To express truncation mutant Env proteins, stop codons were introduced after positions 752, 704, 680 or 592 to produce gp150, gp145, gp140, or gp128, respectively. The Env protein was further changed by deleting amino acids 503 to 537 and 593 to 619, which removes the cleavage site sequence, the fusion domain, and a part of the spacer between the two heptad repeats. The structures of the synthetic HIV envelope genes are shown (Fig. I). The cDNAs were cloned in the expression vector pNGVL or pVR1012 under the control of the cytomegalovirus immediate-early enhancer, promoter, and first intron. The protein sequence was identical to HXB2 Env except for the following amino acid substitutions: F53L, N94D, K192S, I215N, A224T, A346D, P470L, T723I, and S745T.

25

##### **Expression of Envelope Proteins in Transfected Cells**

293 cells ( $10^6$ ) were plated in 60mm dishes. Cells were transfected on the following day with 2  $\mu$ g of plasmid using calcium phosphate (Graham FL, van der Eb AJ, 1973, Virology, 52:456-67). Cells were harvested 48 hours after transfection and lysed in buffer containing 50mM HEPES pH 7.0, 250mM NaCl and 0.5% NP40. The protein concentration in the lysates was determined using the Bradford reagent (Bio-Rad). Proteins (25  $\mu$ g) in lysates were separated by 7.5% SDS-PAGE

30

and transferred to Immobilon-P membrane (Millipore, Bedford, MA). Env was detected by immunoprecipitation followed by Western blotting using polyclonal antibody against gp160 (Intracel, Rockville, MD).

#### **Cell Surface Expression of Envelope by FACS Analysis**

5           293 cells were harvested 48 hours after transfection and washed twice with phosphate-buffered saline (PBS) containing 1% bovine serum albumin and incubated for 30 minutes on ice with polyclonal immunoglobulin from an HIV-1 infected patient. The secondary antibody against human IgG conjugated with FITC (Jackson Immuno Research) was added, incubated for 30 minutes on ice, washed 3  
10 times with PBS, and analyzed by flow cytometry (FACScan). The median fluorescent intensity values were derived using Cell Quant software.

#### **DNA Injection in Mice**

          Six week old, female BALB/c mice were injected intramuscularly with 100 µg of purified plasmid DNA suspended in 200 µl of normal saline. For each  
15 plasmid DNA, a group of 4 mice was injected three times at intervals of two weeks. The mice were bled two weeks after the last injection, sera collected and stored at 4°C.

#### **Quantitation of the Antibody Response**

          Immunoprecipitation and Western blotting was used to detect the antibodies  
20 that bind to native envelope proteins. Sera from immunized mice were used to immunoprecipitate gp160 from cell lysates of gp160-transfected 293 cells. Indicated dilutions were used to immunoprecipitate gp160 from the cell lysate (400 µg). Immunocomplexes were separated by 7.5% SDS-PAGE and analyzed by immunoblotting using the polyclonal antibody against gp160.

#### **Analysis of CTL Response**

          Spleens were removed aseptically and gently homogenized to a single cell suspension, washed, and resuspended to a final concentration of  $5 \times 10^7$  cells/ml. Cells were incubated for 7 days in presence of IL-2 (10 U/ml) and either irradiated peptide-pulsed splenocytes from naive mice or an irradiated stable cell line  
30 expressing the full length gp160 *BC-env/rev*, (Irwin MJ et al., 1994, *J Virol*, 68:5036-44). Three types of target cells were used: peptide-pulsed P815 cells (ATCC TIB64), BC10ME cells stably expressing gp160, and BC10ME cells (Collins JL, Patek PQ, Cohn M. Tumorigenicity and lysis by natural killers. *J Exp Med*, 1981; 153:89-106) pulsed with peptides derived from gp160 sequence. Target



cells were labeled with  $^{51}\text{Cr}$  for 90 minutes and washed three times with RPMI-1640 with 10% FBS, 2mM glutamine,  $5 \times 10^{-5}\text{M}$   $\beta$ -mercaptoethanol and fungisone, 250units/ml, and resuspended in this media. Cytolytic activity was determined in triplicate samples using all different target cell dilutions in a 5-hour  $^{51}\text{Cr}$ -release assay (Ohno T et al., 1997, Gene Ther, 4:361-6).

### *Gag and Pol*

#### **Development of Synthetic HIV-1 Gag-Pol Expression Vectors**

The protein sequences of Gag (amino acids 1-432) from HXB2 (GenBank accession number K03455) and Pol (amino acids 3-1003) from NL4-3 (GenBank accession number M19921) were reverse translated by the GCG Package (Genetic Computer Group, Inc., Madison, WI) using codons expected for human cells. A 226-bp fragment spanning the frame shift site and the overlapping region of the two reading frames from NL4-3 were retained to allow expression of Gag and Gag-Pol precursor polyproteins in the same construct. 86 oligonucleotides covering 4325 DNA base pairs with 5' SalI and 3' EcoRI sites were purchased from GIBCO Life Technologies. Each of the oligonucleotides was 75 base-pairs with 25 nt of overlap. The codon optimized Gag-Pol gene (hGag-Pol) was assembled by PCR with Pwo (Boehringer Mannheim) and Turbo Pfu (Stratagene) high fidelity DNA polymerase. The PCR conditions were optimized with a PCR optimization kit (Stratagene) on a gradient Robocycler (Stratagene). Full-length synthetic Gag-Pol gene was cloned into the Sal I and blunted Bgl II site of the mammalian expression vector pNGVL-3 (Xu, Ling et al., 1998, Nature Medicine, 4:37-42) and confirmed by DNA sequencing. Three additional constructs were derived from the hGag-Pol gene. 5 Thymidines (Ts) in the frame shift site (FS) of the hGag-Pol gene were deleted ( $\Delta\text{FS}$ ) and the protease was inactivated by replacing AGG in protease to GGC (R42G) to create hGag-Pol $\Delta\text{FS}\Delta\text{Pr}$  (Hung, M, 1998, J Virol, 72:4819-4824; Loeb, DD et al., 1989, Nature 340:397-400). 432 amino acids of the NH<sub>2</sub>- terminal of hGag-Pol gene were deleted and an ATG start codon was added to create the hPol gene. 925 amino acids of the COOH-terminal hGag-Pol were deleted to create the hGag gene. hGagPol $\Delta\text{FS}\Delta\text{Pr}$ , hPol and hGag genes were expressed in the pNGVL-3 plasmid, derived by insertion of a polylinker into pVR1012. The plasmid expressing viral Gag-Pol, pCMV $\Delta\text{8.2}$  was a kind gift (Naldini, L et al., 1996, Science, 272:263-267).

### Transient Transfection and Analysis of Expression

293T cells were maintained in Dulbecco's modified Eagle medium (DMEM; GIBCO-BRL), supplemented with 10% fetal bovine serum (FBS). Plasmid DNAs were purified with double cesium chloride sedimentation gradients. Approximately  
5  $3 \times 10^6$  293T cells were placed in a 10-cm dish one day before transfection. 10  $\mu$ g of pCMVdR8.2 plasmid (containing the viral gag-pol gene), or 5  $\mu$ g of pVR1012s (containing the codon altered genes), were used to transfect 293T cells, using the calcium phosphate method (Chen, C and H Okayama, 1987, Mol. Cell. Biol, 7:2745-2752). Three days after transfection, cell lysates were prepared with RIPA buffer  
10 (Boehringer Mannheim, Indianapolis, IN) and separated by 4-15% gradient SDS-polyacrylamide gel electrophoresis (PAGE), then transferred onto an Immobilon P membrane (Millipore). Membranes were then incubated with anti HIV-1-IgG (AIDS Research and Reference Reagent Program), monoclonal anti-p24 (ICN), or rabbit anti-RT (Intracel, Rockville, MD). Bands were visualized using the ECL  
15 Western blotting detection reagent (Amersham Pharmacia Biotech, Piscataway, NJ), as described by the manufacturer. Expression levels were determined using a phosphorimager.

### Generation of Stably Transfected Cell Lines

hGag and hPol genes were individually subcloned into the Xho I and EcoRI  
20 sites of a retroviral vector, pPGS-CITE-Neo. Three plasmid systems were used to produce recombinant retroviruses containing the hGag or hPol genes (Yang, Z et al., 1998, Science, 279:1034-1037). 48 hours after transfection, the supernatants were collected to transduce CT26 and BC10ME (Collins, JL, 1981, J Exp. Med, 153:89-106) which are syngeneic to Balb/C mice, and selected in 0.8 mg/ml of G418 two  
25 days after infection. The positive clones were screened and confirmed by Western blotting and maintained in 10% FCS supplemented RPMI (GIBCO-BRL) with 0.5mg/ml G418.

### DNA Vaccination of Mice

Female BALB/C mice, 6-8 weeks old, were used for immunogenicity  
30 studies. For DNA vaccination, mice were immunized with 100  $\mu$ l ( 0.5  $\mu$ g/ml DNA , 0.9% NaCl) in the quadriceps muscle of each hind leg every two weeks for a total of 4 injections.

**CTL Assay**

Animals were sacrificed after the last immunization and spleens removed from both naïve and immunized mice using aseptic techniques. Splenic lymphocytes were harvested and the chromium release CTL assay was performed in triplicate as previously described (Ohno, T. et al., 1997, Gene Ther, 4:361-366). The peptides used for sensitizing cells are as follows: Two peptide mixtures from the Gag protein P17(88-115) and p24(62-76). Seven peptide mixtures from the Pol protein: 1) P66(175-189), 2) P66(179-193), 3) P66(183-197), 4) P66(187-201), 5) P66(223-237), 6) P66(227-241), 7) P66(367-381).

**Measurement of Antibody Responses***Anti-p24 ELISA assays*

The anti-p24 ELISA assay was performed in Immunlon ninety-six-well plates (Dynet Technologies Inc., Chantilly, VA). The plates were coated with 50 µl of purified recombinant HIV-1<sub>IIIB</sub> p24 antigen (Intracel) at a concentration of 2 µg/ml in PBS buffer, pH 7.4 (GIBCO) with 0.05% sodium azide. Plates were washed 3x in PBS containing 3% BSA and 0.05% Tween20 (blocking buffer) and incubated for 2 hours. Mouse sera were serially diluted from 1:100 to 1:12,800 in blocking buffer, added to the p24-coated plates and incubated overnight at 4°C. Plates were then washed four times with PBS (0.05% Tween 20), and incubated with goat anti-mouse IgG (1:10,000 dilution, Sigma) for 2 hours at room temperature. Plates were washed four times, and then pNPP alkaline phosphatase substrate (75 µl; Sigma, St. Louis, MO) was added to each well. The reaction was stopped after one hour by addition of 0.5 N NaOH (25 µl). The plates were read on an ELISA reader at 405 nm, and titers were calculated at a cutoff optical density of 0.4.

**HIV-1 immunoblotting**

The strips containing HIV-1 proteins (Immunectics Inc., Cambridge, MA) were incubated with pooled mouse sera at a dilution of 1:25. Purified human anti-HIV IgG (AIDS Research and Reference Reagent Program, Rockville, MD) was used as a positive control. Bands were visualized using the ECL Western blotting detection reagent (Amersham Pharmacia Biotech, Piscataway, NJ).

*Immunoprecipitation and Western blotting*

Three days after transfection, hPol gene-transfected 293T cell lysates were prepared with RIPA buffer. The pooled mouse sera were diluted with immunoprecipitation (IP) buffer (100mM KCl, 2.5mM MgCl<sub>2</sub>, 20mM HEPES, pH

7.9, 0.1% NP-40, 1mM DTT and proteinase inhibitors). After adding 10 µg of the cell lysate containing the HIV-1 Pol protein, the reactions were incubated overnight on a rotator at 4°C. 250 µl of Protein G and A Sepharose beads (10%V/V in IP buffer) were then added, and the reactions were incubated for 2 hours on a rotator at 4°C. The beads were washed four times with IP buffer, resuspended in 30 µl of 1X sample buffer and loaded onto SDS-PAGE. After transfer onto an Immobilon P membrane (Millipore), membranes were incubated with anti HIV-1-IgG (AIDS Research and Reference Reagent Program). Bands were visualized using the ECL Western blotting detection reagent (Amersham Pharmacia Biotech).

10

### Prime/Boost Vaccination Strategy

Groups of guinea pigs (4/group) were immunized (primed) with the vectors listed below, 3 times at 2 week intervals. Two weeks after the third DNA immunization, blood was collected for immune analysis. 2-3 days after the blood collection, the animals received a boost with a corresponding adenoviral construct, AdApt (Sullivan, N. et al., 2000, Nature, 408:605-608). Blood was again collected 2 weeks after the adenoviral boost. Blood was tested for neutralizing antibodies and (using an ELISA) the presence of antibody against the envelope. The groups are summarized below.

	Prime	Adenoviral Boost
20	gp140delCFI (R5)	Adv gp140delCFI(R5)
	gp145delCFI (R5)	Adv gp145delCFI(R5)
	gp145delCFI (R5)	Adv gp140delCFI(R5)
	gp140delCFI (89.6P)	Adv gp140delCFI(89.6P)
	gp145delCFI (89.6P)	Adv gp145delCFI(89.6P)
25	gp145delCFI (89.6P)	Adv gp140delCFI(89.6P)
	Adv gp140delCFI (R5)	Adv gp140delCFI(R5)
	Adv gp145delCFI (R5)	Adv gp145delCFI(R5)
	Adv gp145delCFI (R5)	Adv gp140delCFI(R5)
	Adv gp140delCFI (89.6P)	Adv gp140delCFI(89.6P)
30	Adv gp145delCFI (89.6P)	Adv gp145delCFI(89.6P)
	Adv gp145delCFI (89.6P)	Adv gp140delCFI(89.6P)

## Results

5 The level of antibody in the sera after adenoviral boost was 8- to 10-fold higher than the level of antibody present after 3 injections of DNA alone, in almost all of the cases. Neutralizing antibodies were detected in the sera of animals immunized with either gp145delCFI DNA or gp140delCFI DNA followed by Advgp140delCFI.

Neutralization was also observed in animals immunized with gp145delCFI DNA followed by Adv gp145delCFI, and in animals both primed and boosted with adenovirus.

10 Optimal neutralization was seen when animals were primed with gp145delCFI DNA followed by adenovirus expressing gp140delCFI in both clade B (R5-Bal strain) and 89.6P (Clade B, dual tropic strain).

All neutralizing activity was able to be blocked with the V3 peptide of the corresponding strains.

15

\*\*\*\*\*

20 While the present invention has been described in some detail for purposes of clarity and understanding, one skilled in the art will appreciate that various changes in form and detail can be made without departing from the true scope of the invention. All figures, tables, and appendices, as well as patents, applications and publications referred to above are hereby incorporated by reference.



**WHAT IS CLAIMED IS:**

1. A nucleic acid molecule comprising a polynucleotide encoding a modified glycosylation or  $\Delta$ CFI HIV Env, or combination thereof, encoding a plasmid taken from Table 1 (or its modified HIV Env insert), or encoding an analog of said plasmid or insert having at least 95% identity thereto.  
5
2. The nucleic acid molecule of Claim 1 which is plasmid 2100 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
3. The nucleic acid molecule of Claim 1 which is plasmid 2200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
10
4. The nucleic acid molecule of Claim 1 which is plasmid 2300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
5. The nucleic acid molecule of Claim 1 which is plasmid 2302 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
15
6. The nucleic acid molecule of Claim 1 which is plasmid 2400 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
20
7. The nucleic acid molecule of Claim 1 which is plasmid 2700 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
8. The nucleic acid molecule of Claim 1 which is plasmid 2701 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
25
9. The nucleic acid molecule of Claim 1 which is plasmid 2702 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.
10. The nucleic acid molecule of Claim 1 which is plasmid 2706 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
30
11. The nucleic acid molecule of Claim 1 which is plasmid 2707 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.  
35

12. The nucleic acid molecule of Claim 1 which is plasmid 2800 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 13. The nucleic acid molecule of Claim 1 which is plasmid 2801 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

14. The nucleic acid molecule of Claim 1 which is plasmid 2804 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 15. The nucleic acid molecule of Claim 1 which is plasmid 2805 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 16. The nucleic acid molecule of Claim 1 which is plasmid 2810 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

17. The nucleic acid molecule of Claim 1 which is plasmid 2811 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 18. The nucleic acid molecule of Claim 1 which is plasmid 2812 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

19. The nucleic acid molecule of Claim 1 which is plasmid 2813 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 20. The nucleic acid molecule of Claim 1 which is plasmid 2814 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 21. The nucleic acid molecule of Claim 1 which is plasmid 2820 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

22. The nucleic acid molecule of Claim 1 which is plasmid 2821 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

23. The nucleic acid molecule of Claim 1 which is plasmid 2822 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 24. The nucleic acid molecule of Claim 1 which is plasmid 2823 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25. The nucleic acid molecule of Claim 1 which is plasmid 2824 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 26. The nucleic acid molecule of Claim 1 which is plasmid 2825 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 27. The nucleic acid molecule of Claim 1 which is plasmid 2826 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

28. The nucleic acid molecule of Claim 1 which is plasmid 2827 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 29. The nucleic acid molecule of Claim 1 which is plasmid 2828 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30. The nucleic acid molecule of Claim 1 which is plasmid 2829 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 31. The nucleic acid molecule of Claim 1 which is plasmid 2830 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 32. The nucleic acid molecule of Claim 1 which is plasmid 2831 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

33. The nucleic acid molecule of Claim 1 which is plasmid 2832 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

34. The nucleic acid molecule of Claim 1 which is plasmid 2833 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 35. The nucleic acid molecule of Claim 1 which is plasmid 2834 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

36. The nucleic acid molecule of Claim 1 which is plasmid 2835 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 37. The nucleic acid molecule of Claim 1 which is plasmid 2836 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 38. The nucleic acid molecule of Claim 1 which is plasmid 2837 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

39. The nucleic acid molecule of Claim 1 which is plasmid 2838 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 40. The nucleic acid molecule of Claim 1 which is plasmid 2839 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

41. The nucleic acid molecule of Claim 1 which is plasmid 2840 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 42. The nucleic acid molecule of Claim 1 which is plasmid 2841 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 43. The nucleic acid molecule of Claim 1 which is plasmid 2842 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

44. The nucleic acid molecule of Claim 1 which is plasmid 2843 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

45. The nucleic acid molecule of Claim 1 which is plasmid 2844 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 46. The nucleic acid molecule of Claim 1 which is plasmid 2845 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

47. The nucleic acid molecule of Claim 1 which is plasmid 2846 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 48. The nucleic acid molecule of Claim 1 which is plasmid 2847 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 49. The nucleic acid molecule of Claim 1 which is plasmid 2848 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

50. The nucleic acid molecule of Claim 1 which is plasmid 2849 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 51. The nucleic acid molecule of Claim 1 which is plasmid 2850 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

52. The nucleic acid molecule of Claim 1 which is plasmid 2851 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 53. The nucleic acid molecule of Claim 1 which is plasmid 2852 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 54. The nucleic acid molecule of Claim 1 which is plasmid 2853 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

55. The nucleic acid molecule of Claim 1 which is plasmid 2854 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.



56. The nucleic acid molecule of Claim 1 which is plasmid 2860 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 57. The nucleic acid molecule of Claim 1 which is plasmid 2861 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

58. The nucleic acid molecule of Claim 1 which is plasmid 2862 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 59. The nucleic acid molecule of Claim 1 which is plasmid 2863 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 60. The nucleic acid molecule of Claim 1 which is plasmid 2864 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

61. The nucleic acid molecule of Claim 1 which is plasmid 2865 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 62. The nucleic acid molecule of Claim 1 which is plasmid 2866 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

63. The nucleic acid molecule of Claim 1 which is plasmid 2867 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 64. The nucleic acid molecule of Claim 1 which is plasmid 2868 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 65. The nucleic acid molecule of Claim 1 which is plasmid 2869 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

66. The nucleic acid molecule of Claim 1 which is plasmid 2870 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

67. The nucleic acid molecule of Claim 1 which is plasmid 2871 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 68. The nucleic acid molecule of Claim 1 which is plasmid 2872 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

69. The nucleic acid molecule of Claim 1 which is plasmid 2873 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 70. The nucleic acid molecule of Claim 1 which is plasmid 2874 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 71. The nucleic acid molecule of Claim 1 which is plasmid 2900 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

72. The nucleic acid molecule of Claim 1 which is plasmid 3000 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 73. The nucleic acid molecule of Claim 1 which is plasmid 3200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

74. The nucleic acid molecule of Claim 1 which is plasmid 3201 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 75. The nucleic acid molecule of Claim 1 which is plasmid 3202 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 76. The nucleic acid molecule of Claim 1 which is plasmid 3203 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

77. The nucleic acid molecule of Claim 1 which is plasmid 3300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

78. The nucleic acid molecule of Claim 1 which is plasmid 3301 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 79. The nucleic acid molecule of Claim 1 which is plasmid 3400 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

80. The nucleic acid molecule of Claim 1 which is plasmid 3401 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 81. The nucleic acid molecule of Claim 1 which is plasmid 3500 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 82. The nucleic acid molecule of Claim 1 which is plasmid 3600 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

83. The nucleic acid molecule of Claim 1 which is plasmid 3700 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 84. The nucleic acid molecule of Claim 1 which is plasmid 3800 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

85. The nucleic acid molecule of Claim 1 which is plasmid 5200 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 86. The nucleic acid molecule of Claim 1 which is plasmid 5201 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 87. The nucleic acid molecule of Claim 1 which is plasmid 5202 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

88. The nucleic acid molecule of Claim 1 which is plasmid 5203 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

89. The nucleic acid molecule of Claim 1 which is plasmid 5300 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 90. The nucleic acid molecule of Claim 1 which is plasmid 5301 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

91. The nucleic acid molecule of Claim 1 which is plasmid 5303 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 92. The nucleic acid molecule of Claim 1 which is plasmid 5304 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 93. The nucleic acid molecule of Claim 1 which is plasmid 5305 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

94. The nucleic acid molecule of Claim 1 which is plasmid 5306 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 95. The nucleic acid molecule of Claim 1 which is plasmid 5307 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

96. The nucleic acid molecule of Claim 1 which is plasmid 5308 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 97. The nucleic acid molecule of Claim 1 which is plasmid 5309 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 98. The nucleic acid molecule of Claim 1 which is plasmid 5350 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

99. The nucleic acid molecule of Claim 1 which is plasmid 5351 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

100. The nucleic acid molecule of Claim 1 which is plasmid 5352 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 101. The nucleic acid molecule of Claim 1 which is plasmid 5353 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

102. The nucleic acid molecule of Claim 1 which is plasmid 5354 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 103. The nucleic acid molecule of Claim 1 which is plasmid 5355 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 104. The nucleic acid molecule of Claim 1 which is plasmid 5356 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

105. The nucleic acid molecule of Claim 1 which is plasmid 5357 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 106. The nucleic acid molecule of Claim 1 which is plasmid 5358 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

107. The nucleic acid molecule of Claim 1 which is plasmid 5359 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 108. The nucleic acid molecule of Claim 1 which is plasmid 5360 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 109. The nucleic acid molecule of Claim 1 which is plasmid 5361 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

110. The nucleic acid molecule of Claim 1 which is plasmid 5362 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.



111. The nucleic acid molecule of Claim 1 which is plasmid 5363 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 112. The nucleic acid molecule of Claim 1 which is plasmid 5364 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

113. The nucleic acid molecule of Claim 1 which is plasmid 5365 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 114. The nucleic acid molecule of Claim 1 which is plasmid 5366 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 115. The nucleic acid molecule of Claim 1 which is plasmid 5367 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

116. The nucleic acid molecule of Claim 1 which is plasmid 5368 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 117. The nucleic acid molecule of Claim 1 which is plasmid 5369 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

118. The nucleic acid molecule of Claim 1 which is plasmid 5370 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 119. The nucleic acid molecule of Claim 1 which is plasmid 5371 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 120. The nucleic acid molecule of Claim 1 which is plasmid 5372 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

121. The nucleic acid molecule of Claim 1 which is plasmid 5373 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

122. The nucleic acid molecule of Claim 1 which is plasmid 5374 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 123. The nucleic acid molecule of Claim 1 which is plasmid 5375 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

124. The nucleic acid molecule of Claim 1 which is plasmid 5376 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 125. The nucleic acid molecule of Claim 1 which is plasmid 5377 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 126. The nucleic acid molecule of Claim 1 which is plasmid 5378 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

127. The nucleic acid molecule of Claim 1 which is plasmid 5379 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 128. The nucleic acid molecule of Claim 1 which is plasmid 5500 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

129. The nucleic acid molecule of Claim 1 which is plasmid 5501 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 130. The nucleic acid molecule of Claim 1 which is plasmid 5502 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 131. The nucleic acid molecule of Claim 1 which is plasmid 5503 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

132. The nucleic acid molecule of Claim 1 which is plasmid 5504 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

133. The nucleic acid molecule of Claim 1 which is plasmid 5505 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 134. The nucleic acid molecule of Claim 1 which is plasmid 5506 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

135. The nucleic acid molecule of Claim 1 which is plasmid 5507 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 136. The nucleic acid molecule of Claim 1 which is plasmid 5508 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 137. The nucleic acid molecule of Claim 1 which is plasmid 5509 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

138. The nucleic acid molecule of Claim 1 which is plasmid 5510 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 139. The nucleic acid molecule of Claim 1 which is plasmid 5511 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

140. The nucleic acid molecule of Claim 1 which is plasmid 5512 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 141. The nucleic acid molecule of Claim 1 which is plasmid 5513 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 142. The nucleic acid molecule of Claim 1 which is plasmid 5514 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

143. The nucleic acid molecule of Claim 1 which is plasmid 5515 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

144. The nucleic acid molecule of Claim 1 which is plasmid 5516 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 145. The nucleic acid molecule of Claim 1 which is plasmid 5517 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

146. The nucleic acid molecule of Claim 1 which is plasmid 5518 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 147. The nucleic acid molecule of Claim 1 which is plasmid 5519 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 148. The nucleic acid molecule of Claim 1 which is plasmid 5520 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

149. The nucleic acid molecule of Claim 1 which is plasmid 5521 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 150. The nucleic acid molecule of Claim 1 which is plasmid 5522 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

151. The nucleic acid molecule of Claim 1 which is plasmid 5523 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 152. The nucleic acid molecule of Claim 1 which is plasmid 5524 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 153. The nucleic acid molecule of Claim 1 which is plasmid 5525 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

154. The nucleic acid molecule of Claim 1 which is plasmid 5526 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

155. The nucleic acid molecule of Claim 1 which is plasmid 5527 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 156. The nucleic acid molecule of Claim 1 which is plasmid 5528 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

157. The nucleic acid molecule of Claim 1 which is plasmid 5529 (or its modified HIV Env insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 158. A nucleic acid molecule comprising a polynucleotide encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, encoding a plasmid taken from Table 1 (or its modified HIV Gag, Pol, or Nef insert), or encoding an analog of said plasmid or insert having at least 95% identity thereto.

15 159. The nucleic acid molecule of Claim 158 which is plasmid 3900 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 160. The nucleic acid molecule of Claim 158 which is plasmid 3901 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

161. The nucleic acid molecule of Claim 158 which is plasmid 4000 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 162. The nucleic acid molecule of Claim 158 which is plasmid 4001 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

163. The nucleic acid molecule of Claim 158 which is plasmid 4100 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 164. The nucleic acid molecule of Claim 158 which is plasmid 4101 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

35 165. The nucleic acid molecule of Claim 158 which is plasmid 4200 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.



166. The nucleic acid molecule of Claim 158 which is plasmid 4300 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 167. The nucleic acid molecule of Claim 158 which is plasmid 4301 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

168. The nucleic acid molecule of Claim 158 which is plasmid 4302 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

10 169. The nucleic acid molecule of Claim 158 which is plasmid 4303 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

15 170. The nucleic acid molecule of Claim 158 which is plasmid 4304 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

171. The nucleic acid molecule of Claim 158 which is plasmid 4305 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

20 172. The nucleic acid molecule of Claim 158 which is plasmid 4306 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

173. The nucleic acid molecule of Claim 158 which is plasmid 4308 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

25 174. The nucleic acid molecule of Claim 158 which is plasmid 4309 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

30 175. The nucleic acid molecule of Claim 158 which is plasmid 4310 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

176. The nucleic acid molecule of Claim 158 which is plasmid 4311 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

177. The nucleic acid molecule of Claim 158 which is plasmid 4312 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

5 178. The nucleic acid molecule of Claim 158 which is plasmid 4313 (or its modified HIV Gag, Pol, or Nef insert) or an analog of said plasmid or insert having at least 95% identity thereto.

179. A modified HIV Env, Gag, Pol, or Nef encoded by the nucleic acid molecule of any of Claims 1-178.

10 180. The nucleic acid molecule of any of Claims 1-178 further comprising a backbone, wherein said backbone is a plasmid vector.

181. The nucleic acid molecule of any of Claims 1-178 further comprising a backbone, wherein said backbone is an adenoviral vector.

15 182. A nucleic acid molecule comprising a polynucleotide encoding a modified glycosylation or  $\Delta$ CFI HIV Env, or combination thereof, encoding a modified HIV Env insert taken from a plasmid taken from Table 1, or encoding an analog of said insert having at least 95% identity thereto, and further encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, encoding a modified HIV Gag, Pol, or Nef insert taken from a plasmid taken from Table 1, or encoding an analog of said insert having at least 95% identity thereto.

20 183. A pharmaceutical composition comprising at least one nucleic acid molecule of any of Claims 1-178 and a pharmaceutically acceptable solution in a therapeutically effective dose

25 184. A vaccine composition comprising at least one nucleic acid molecule of any of Claims 1-178 and a pharmaceutically acceptable solution in a prophylactically effective dose.

185. A pharmaceutical composition comprising at least one modified HIV Env, Gag, Pol, or Nef of Claim 179 and a pharmaceutically acceptable solution in a therapeutically effective dose.

30 186. A vaccine composition comprising at least one modified HIV Env, Gag, Pol, or Nef of Claim 179 and a pharmaceutically acceptable solution in a prophylactically effective dose.

187. The composition of any of Claims 183-186 additionally comprising an adjuvant.

35 188. The composition of Claim 187 wherein said adjuvant is a cytokine.

189. The composition of Claim 183 or 184 which combines plasmid 4306 (gag-pol-nef) and plasmid 2827 (env).

190. The composition of Claim 183 or 184 which combines plasmid 4306 (gag-pol-nef Clade B), plasmid 2827 (env Clade B), plasmid 4311 (gag-pol-nef Clade C), and plasmid 5509 (env Clade C).

191. A method comprising the step of administering by genetic immunization a nucleic acid molecule of any of Claims 1-178 to a host to ameliorate the symptoms of AIDS or HIV infection.

192. A method comprising the step of administering by genetic immunization a nucleic acid molecule of any of Claims 1-178 to a host to generate an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.

193. A method comprising the step of administering by protein immunization a modified HIV Env, Gag, Pol, or Nef of Claim 179 to a host to ameliorate the symptoms of AIDS or HIV infection.

194. A method comprising the step of administering by protein immunization a modified HIV Env, Gag, Pol, or Nef of Claim 179 to a host to generate an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.

195. The method of any of Claims 191-194 wherein said step is a primary immunization.

196. The method of any of Claims 191-194 wherein said step is a boost.

197. The method of Claims 191 or 192 wherein said step is a primary immunization with a plasmid vector.

198. The method of Claims 191 or 192 wherein said step is a boost with an adenoviral vector.

199. The method of Claims 191 or 192 wherein said step is a primary immunization with a plasmid vector expressing gp145delCFI and further comprising a step which is a boost with an adenoviral vector expressing gp140delCFI.

200. A nucleic acid molecule of any of Claims 1-178 for use by genetic immunization in amelioration of the symptoms of AIDS or HIV infection.

201. A nucleic acid molecule of any of Claims 1-178 for use by genetic immunization in generation of an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.

202. A modified HIV Env, Gag, Pol, or Nef of Claim 179 for use by protein immunization in amelioration of the symptoms of AIDS or HIV infection.

203. A modified HIV Env, Gag, Pol, or Nef of Claim 179 for use by protein immunization in generation of an antibody or CTL response against native HIV Env, Gag, Pol, or Nef.

5 204. A nucleic acid molecule comprising a polynucleotide encoding a modified  $\Delta$ CFI HIV Env or encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, or a combination thereof, applied to any HIV clade or strain.

10 205. A nucleic acid molecule comprising a polynucleotide encoding a modified  $\Delta$ CFI HIV Env or encoding in a continuous open reading frame a modified HIV Gag-Pol or Gag-Pol-Nef fusion protein, or a combination thereof, applied to related proteins of other viruses.

15 206. A nucleic acid molecule comprising a polynucleotide encoding a modified viral membrane fusion protein with deletions in sufficient domains to expose a core protein for optimal antigen presentation, wherein the viral membrane fusion protein is defined by a central triple stranded coiled coil followed by a disulfide-bonded loop that reverses the chain direction and connects to an  $\alpha$  helix packed antiparallel to the core helices.

20 207. A nucleic acid molecule comprising a polynucleotide encoding in a continuous open reading frame a modified viral protein, wherein the viral protein is transcribed in native form by translational frame shifting in a non-equivalent ratio or through a separate open reading frame.

Figure A

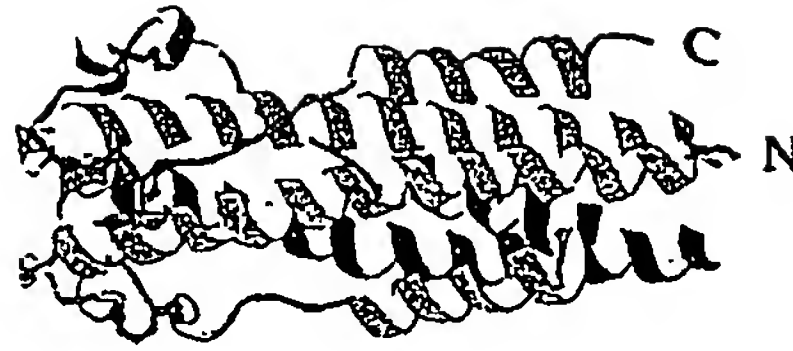
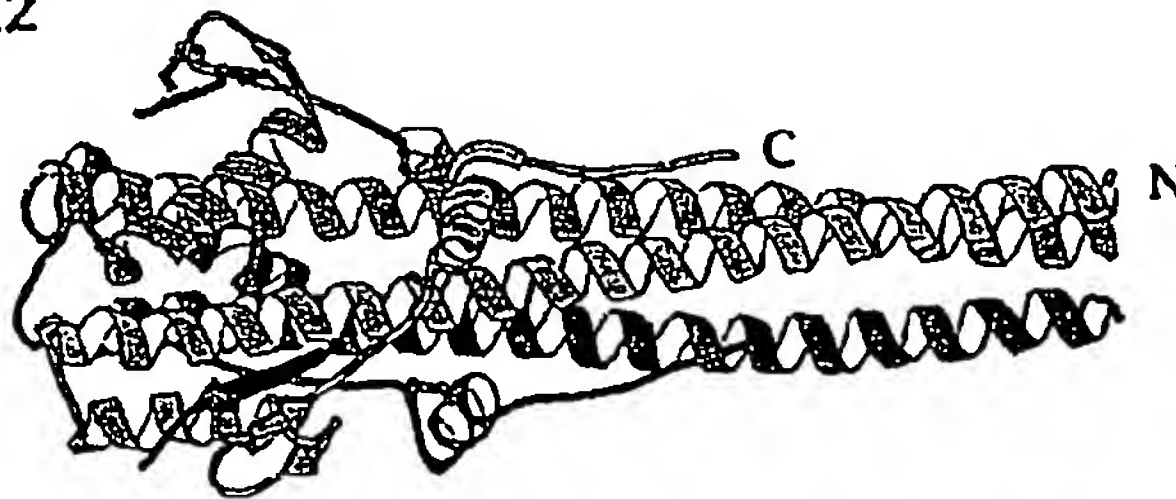
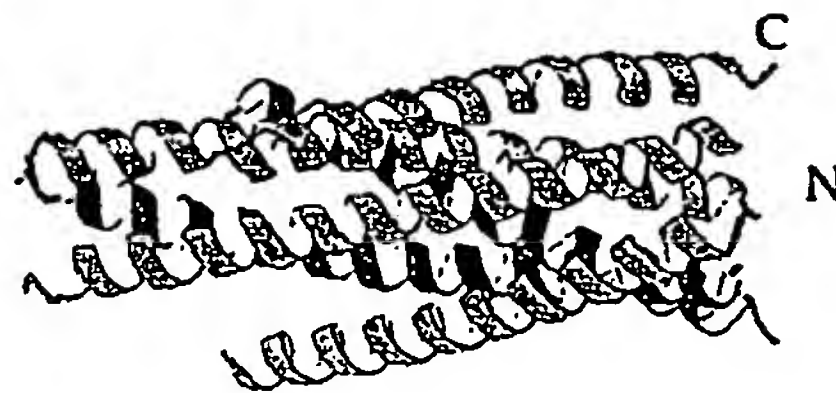
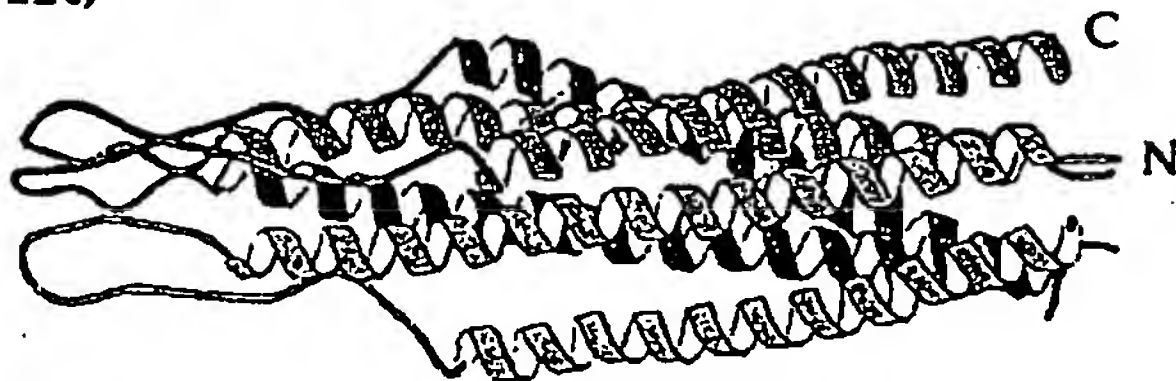
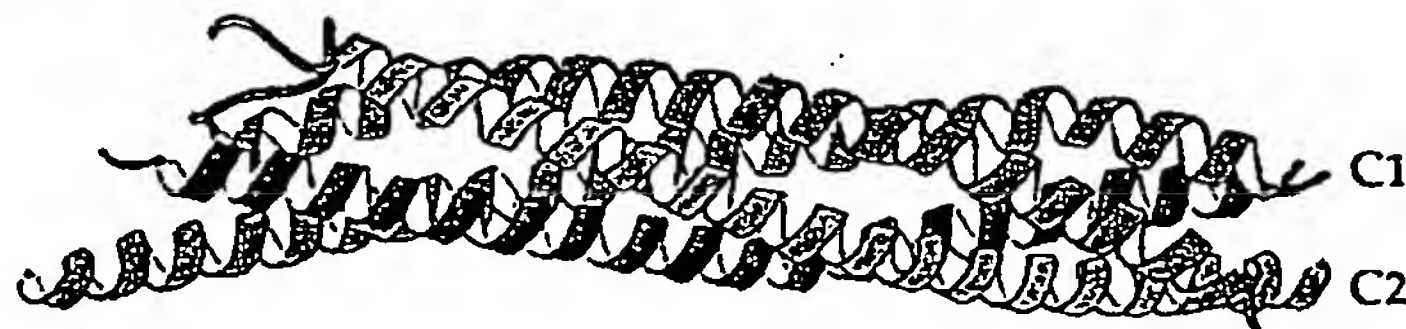
**A Ebola GP2****B MoMuLv Mo-55****C low-pH-treated  
influenza HA2****D protease resistant  
core of HIV gp41****E SIV gp41 (NMR)****F SNAREs**



Figure I

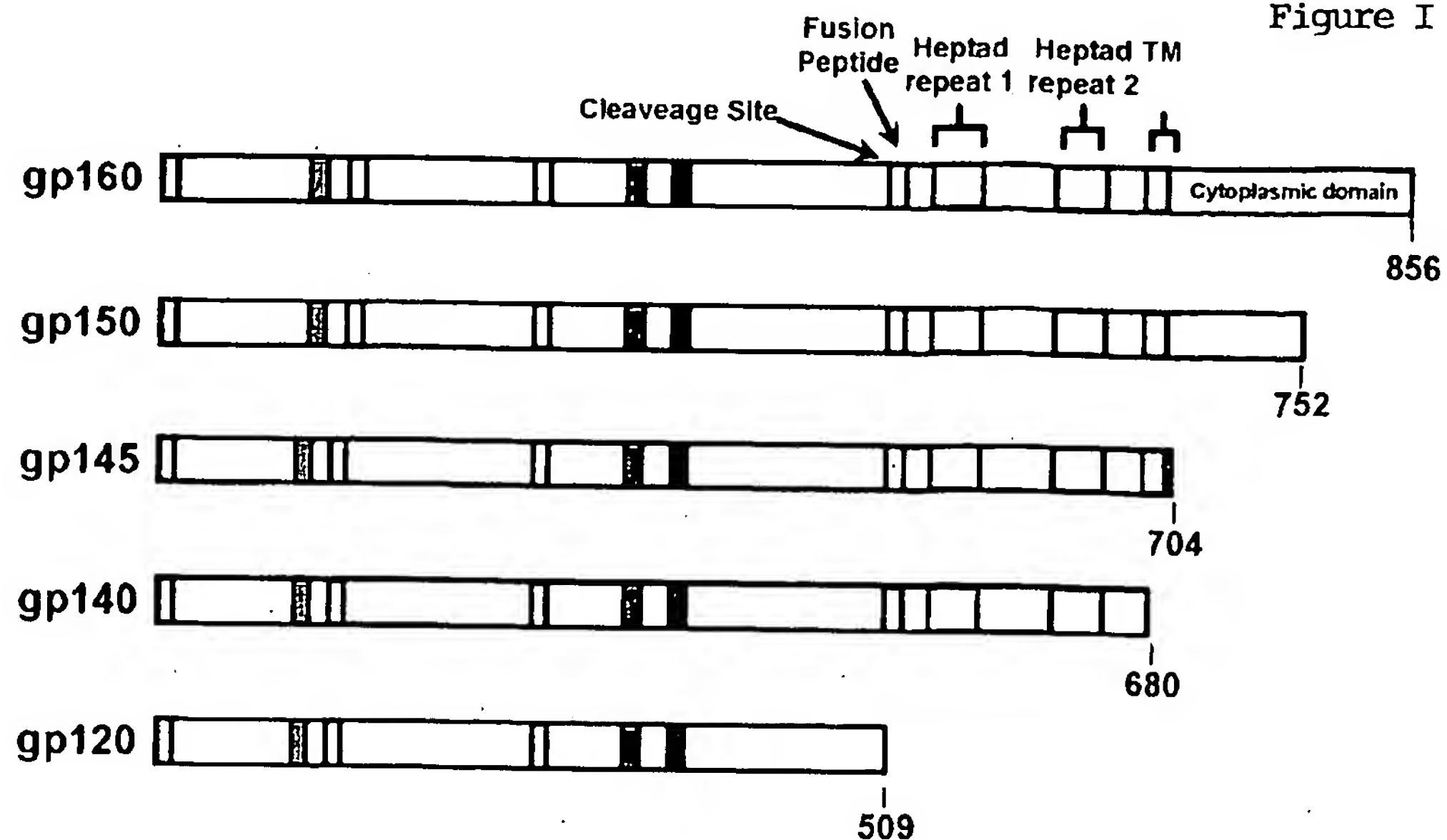
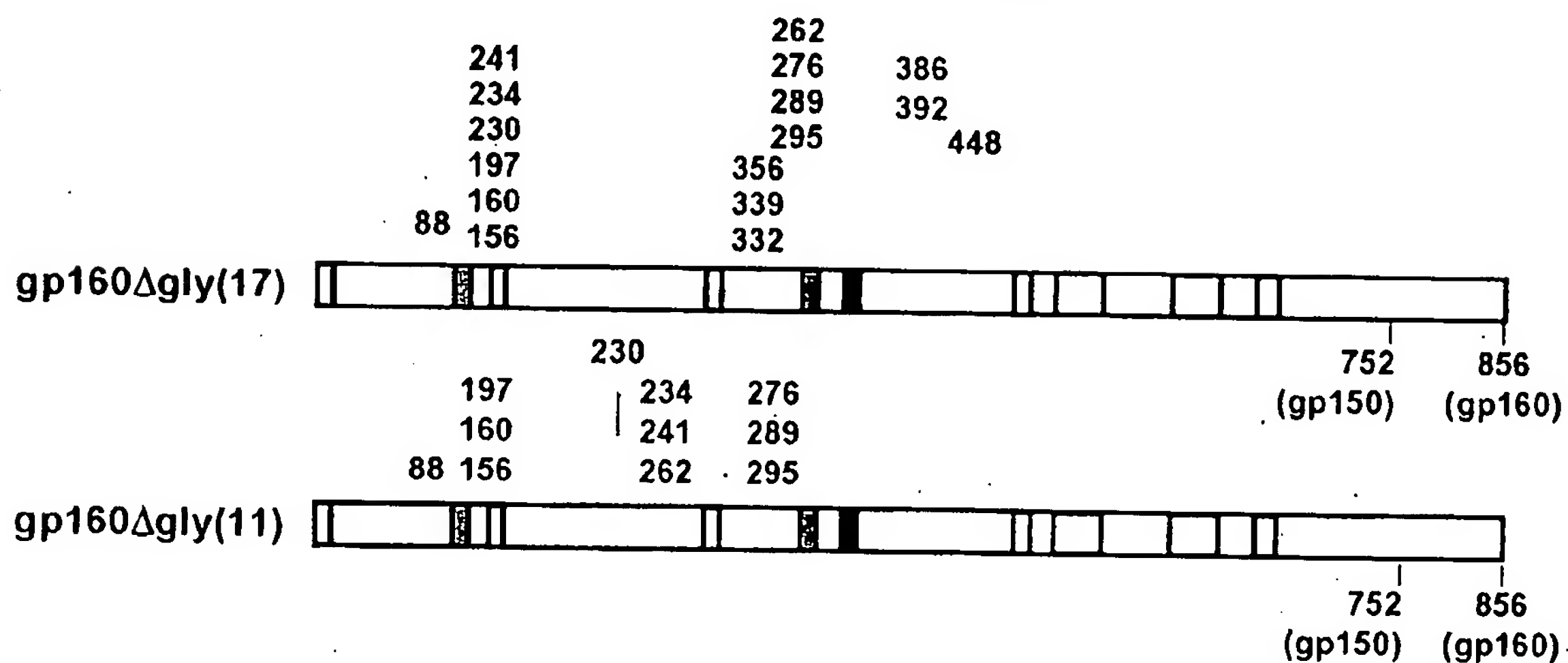
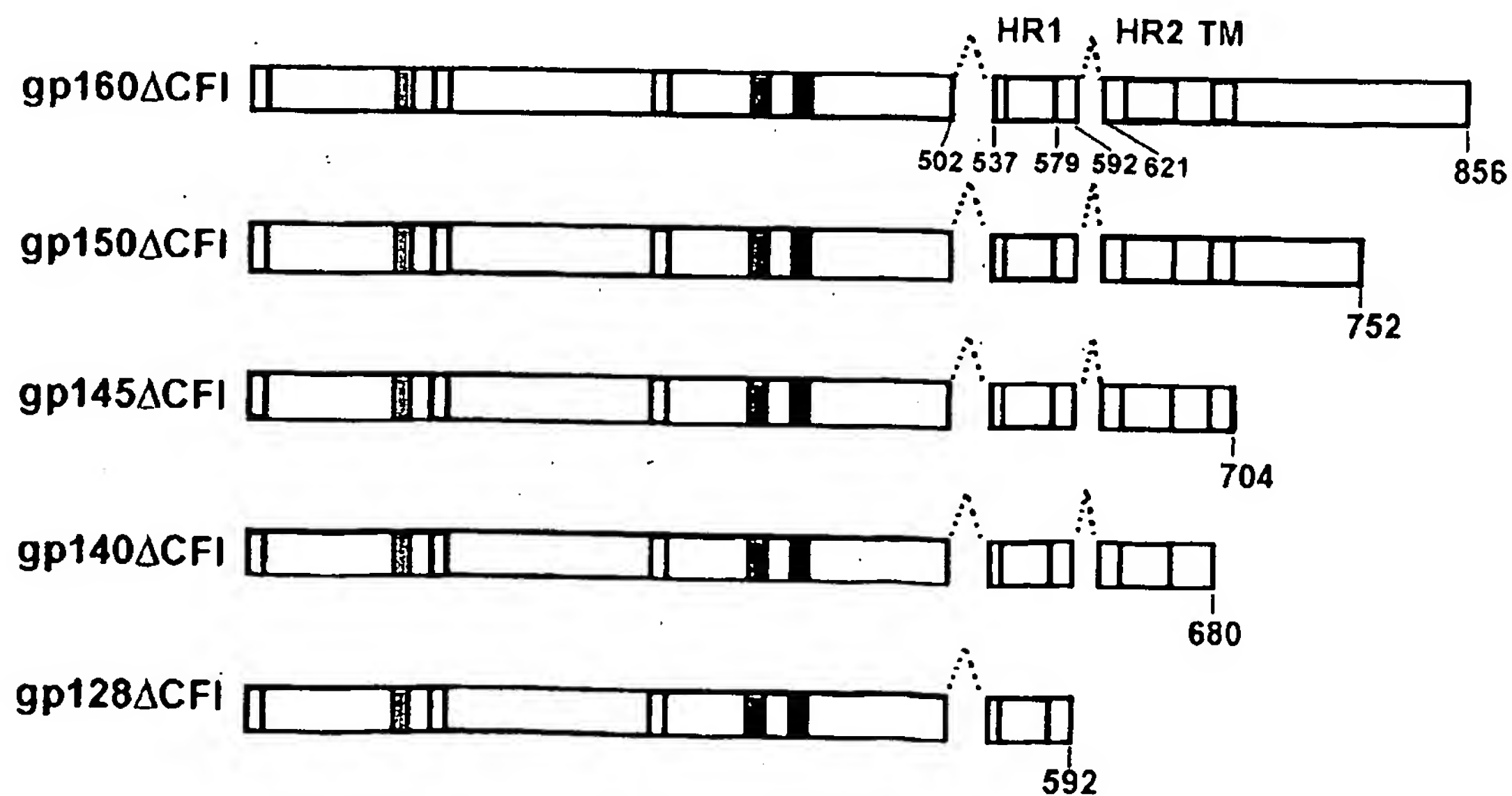
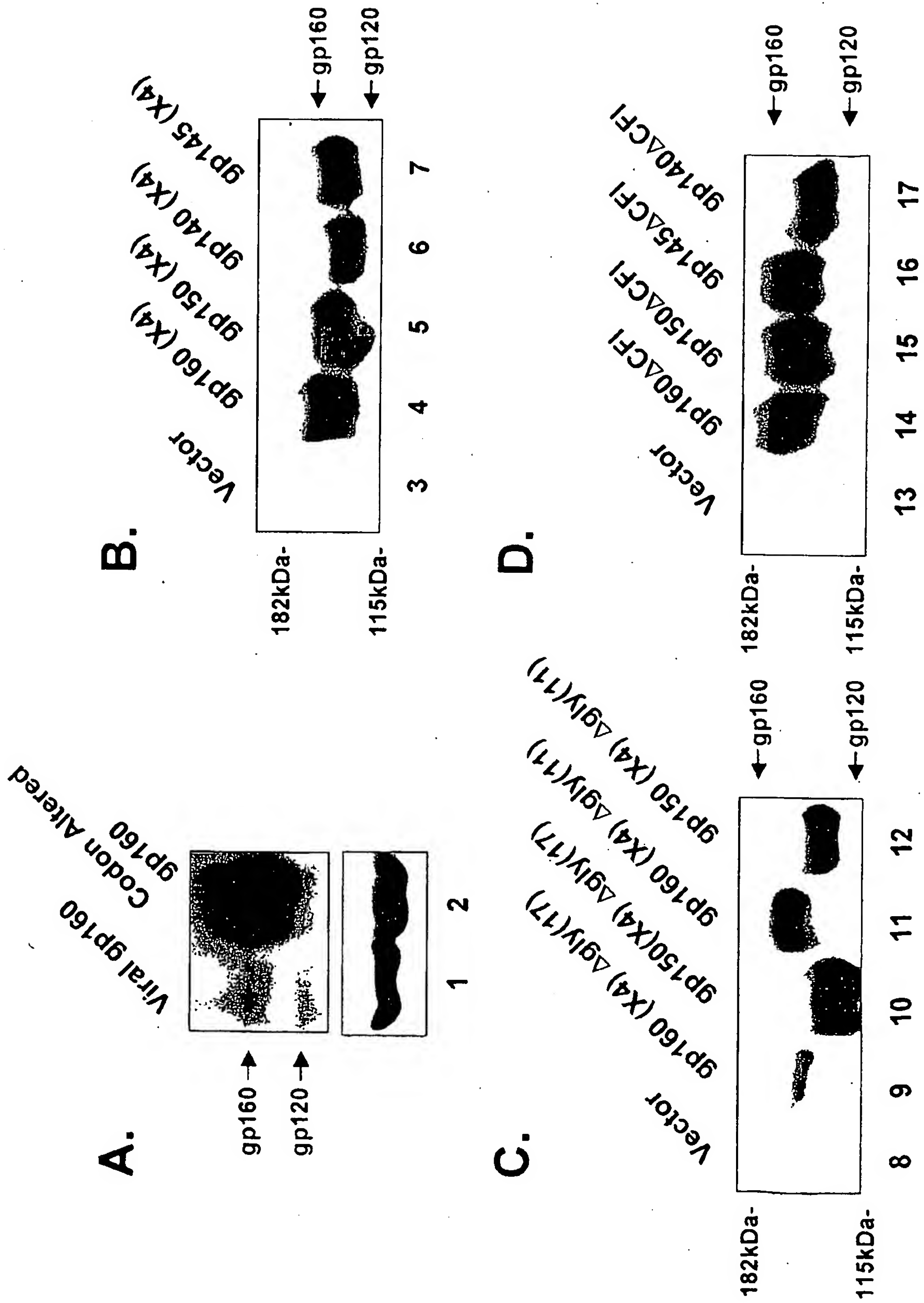
**A.****B.****C.**

Figure II



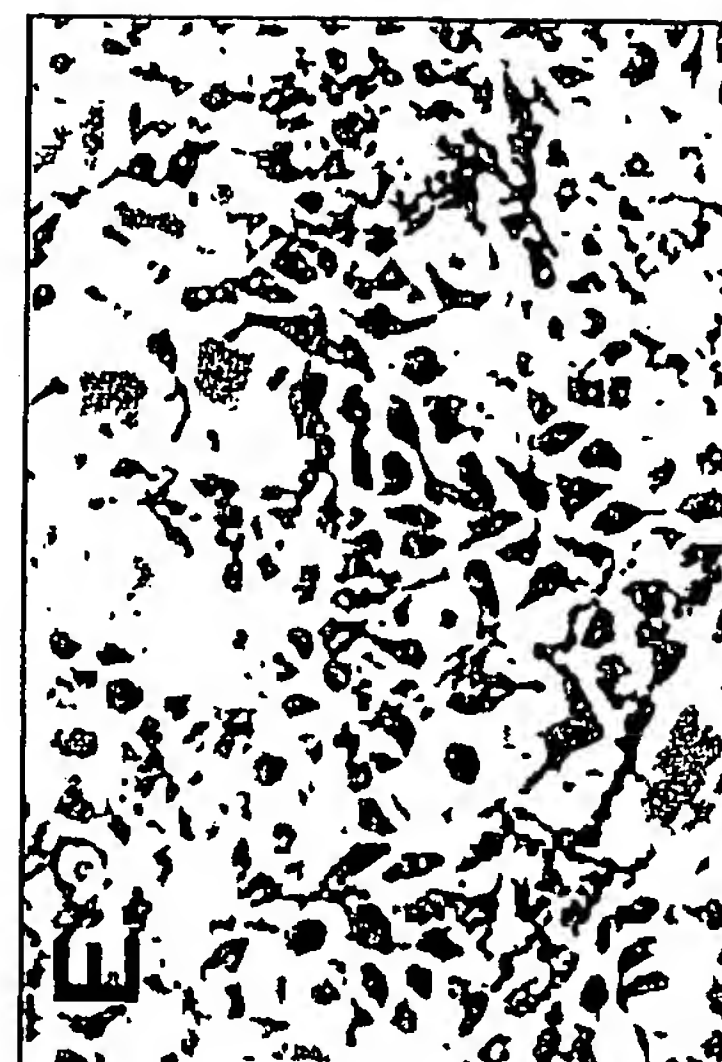
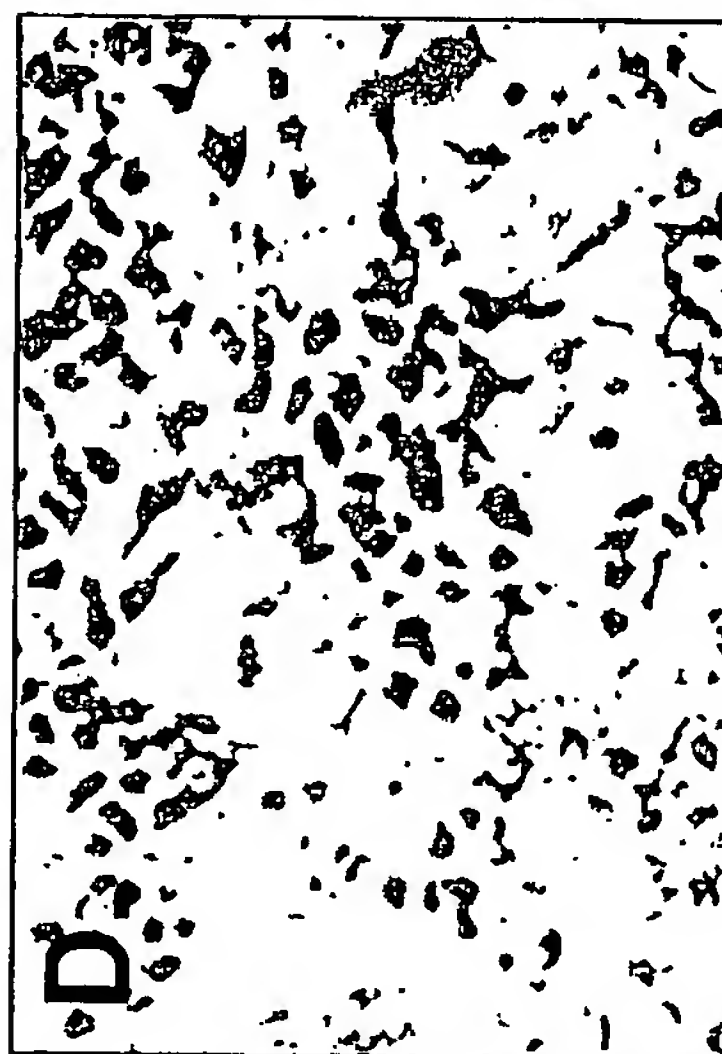


Figure IV

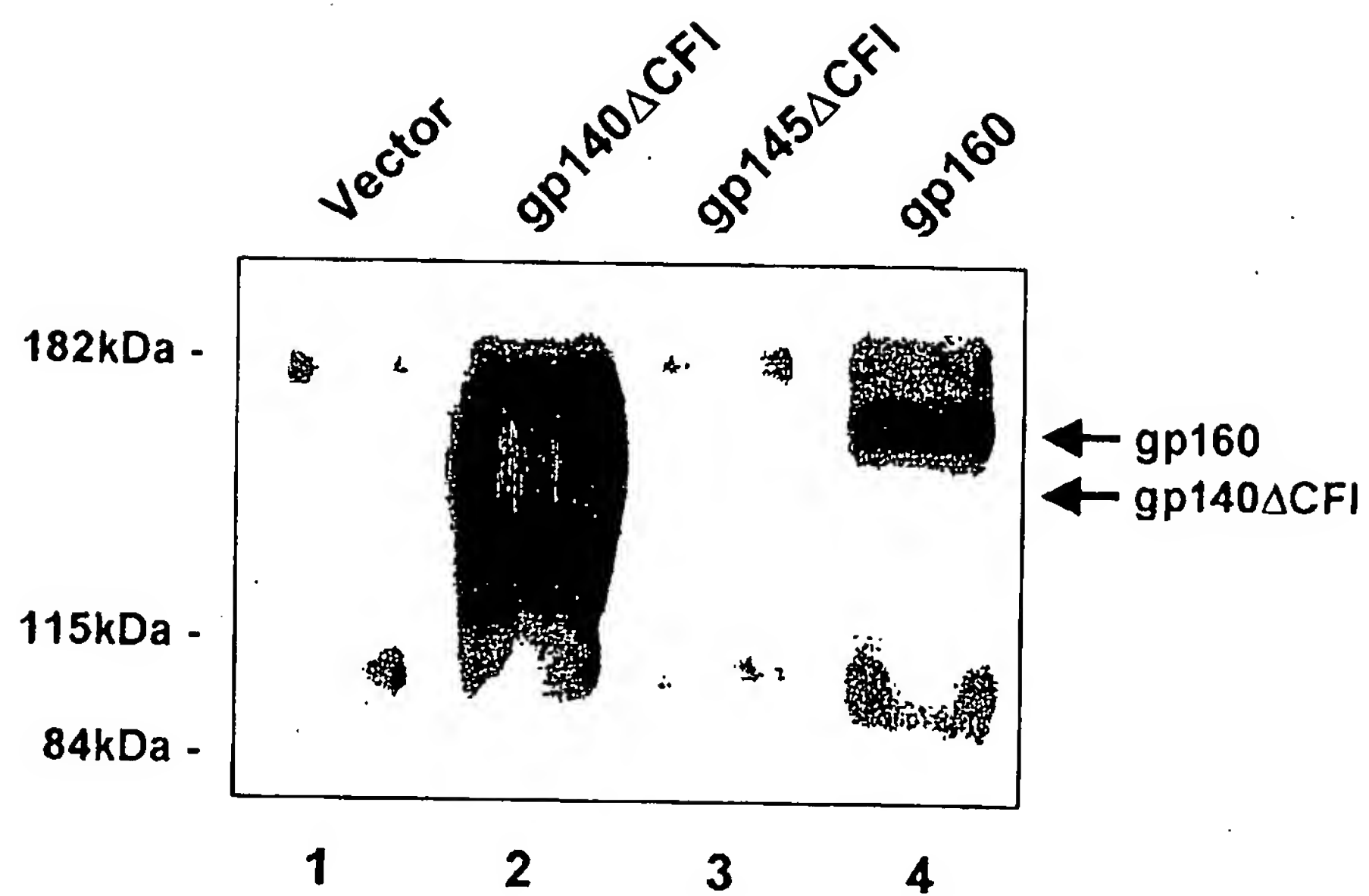
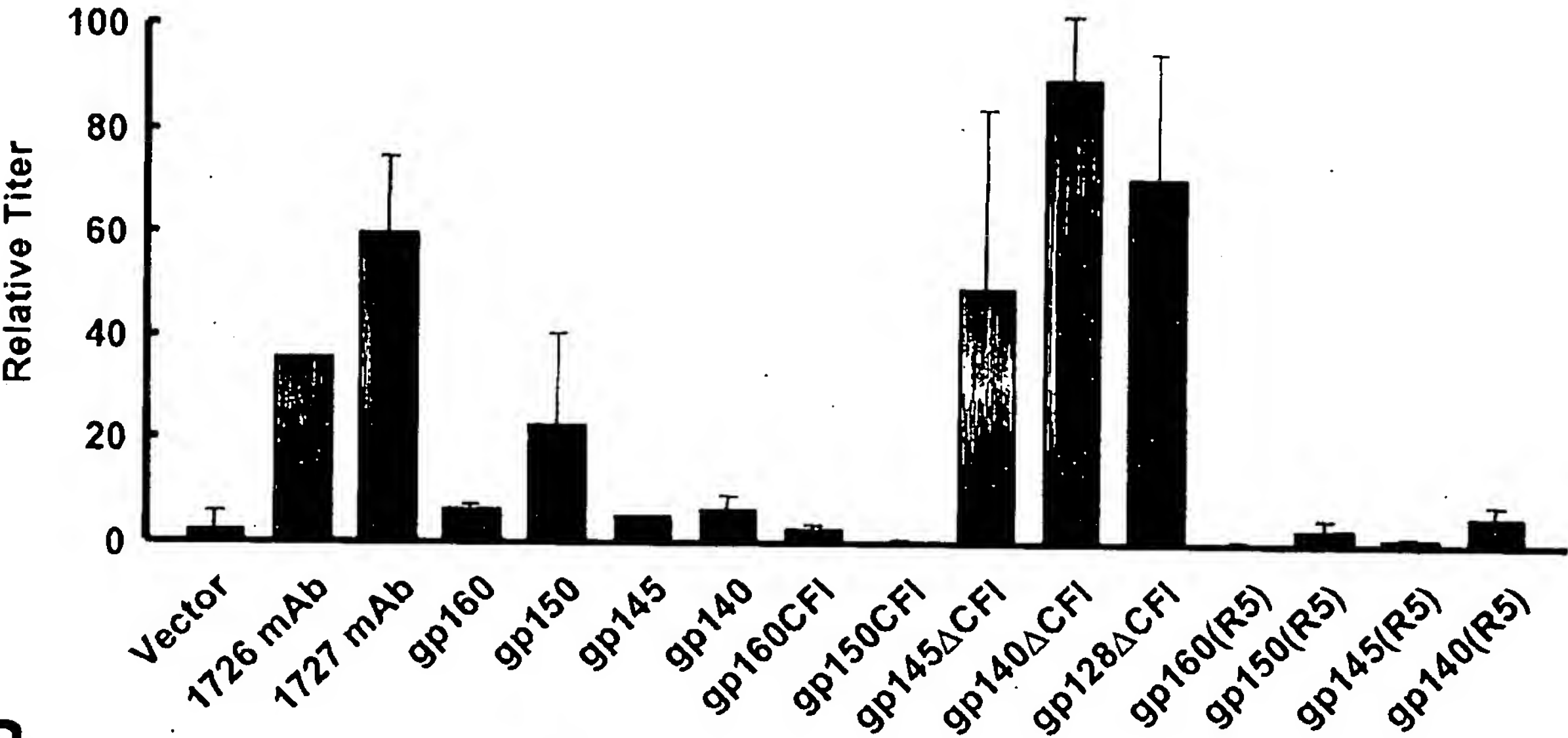
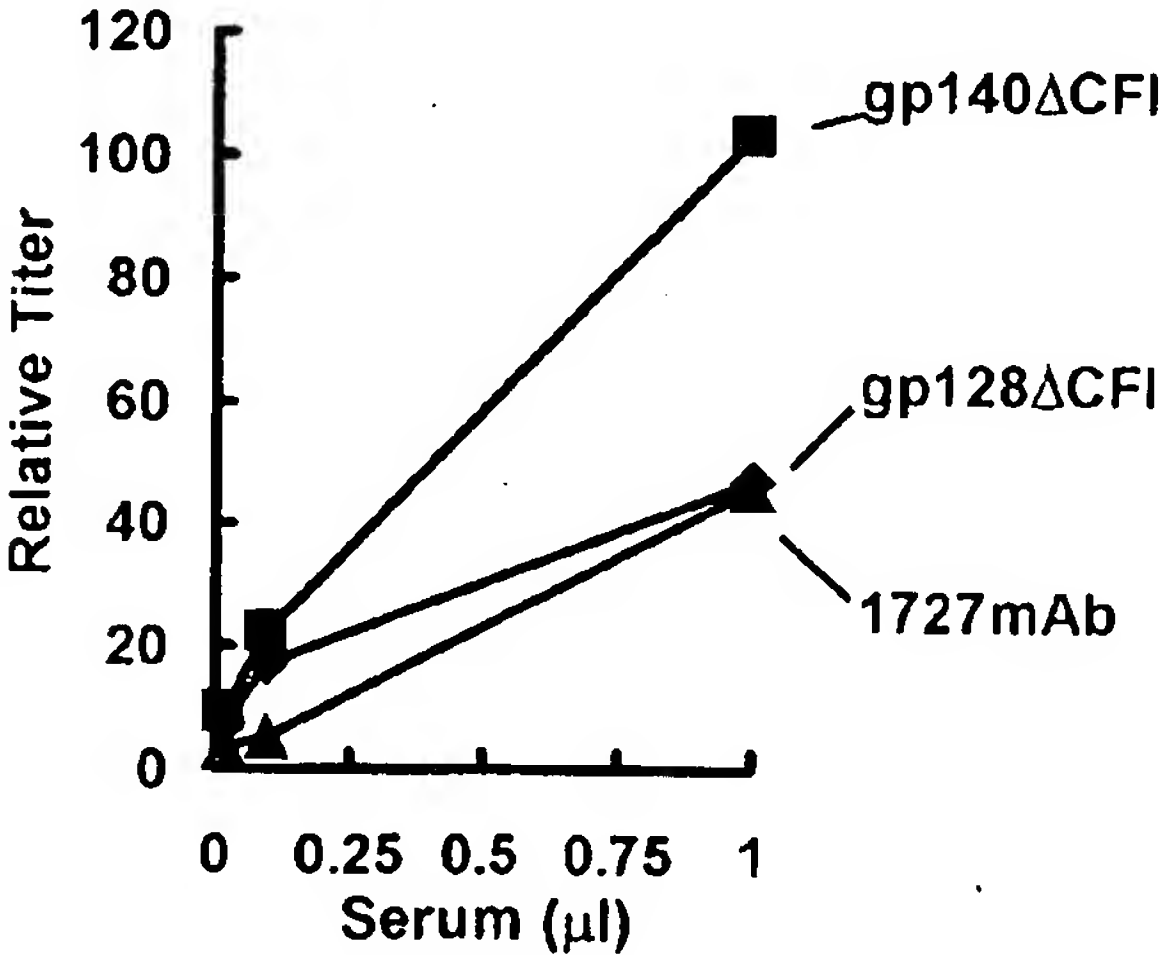


Figure V

A.



B.



C.

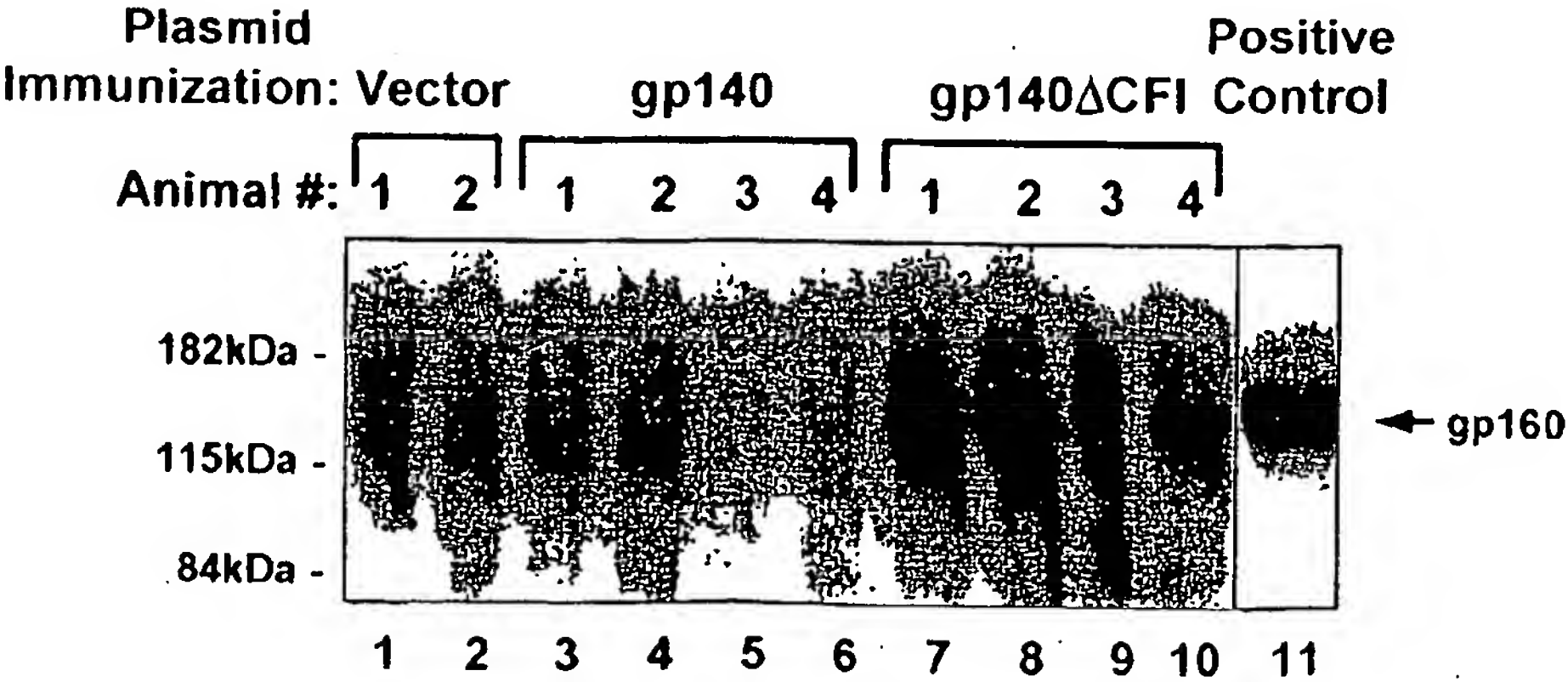




Figure VI

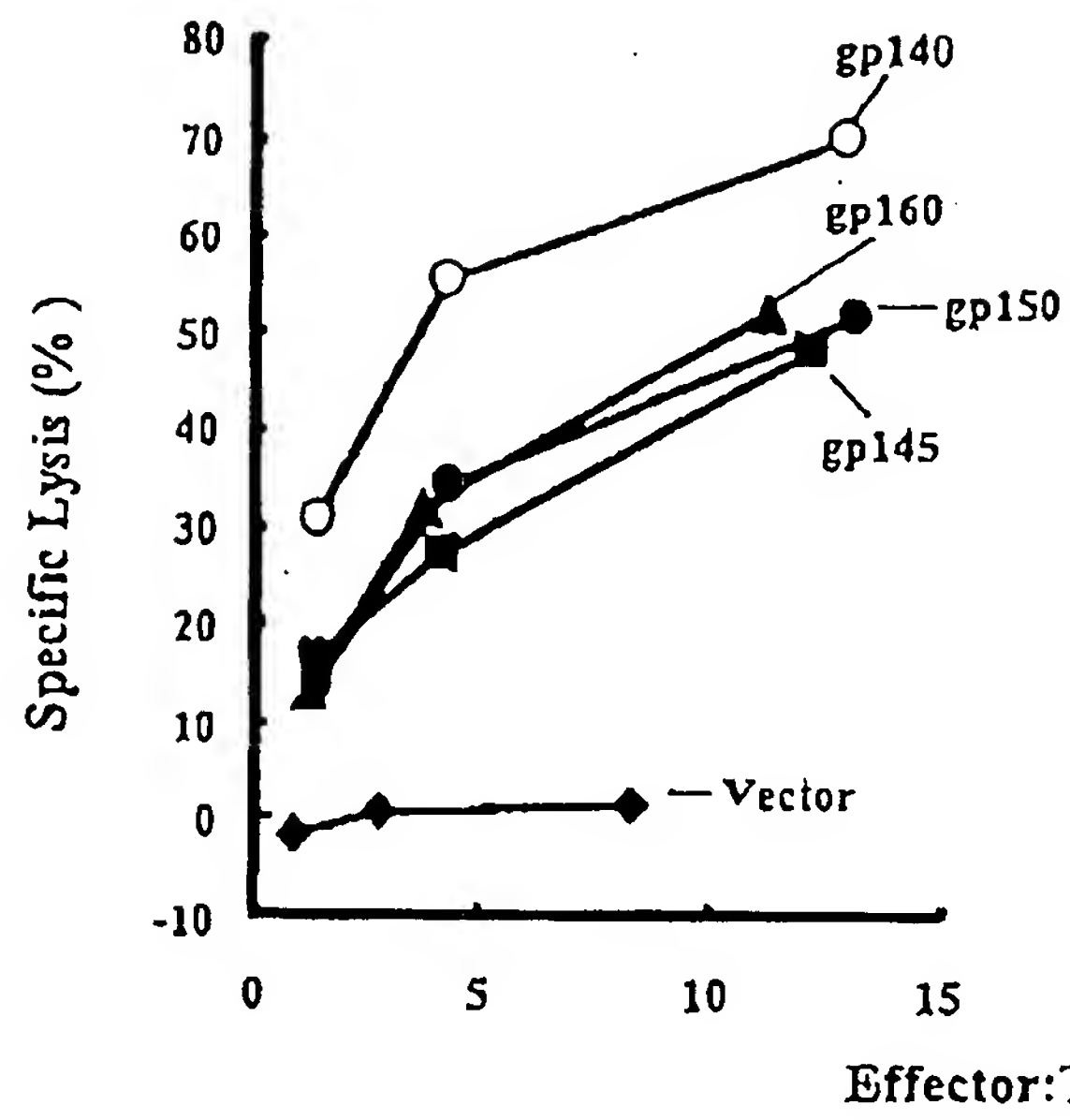
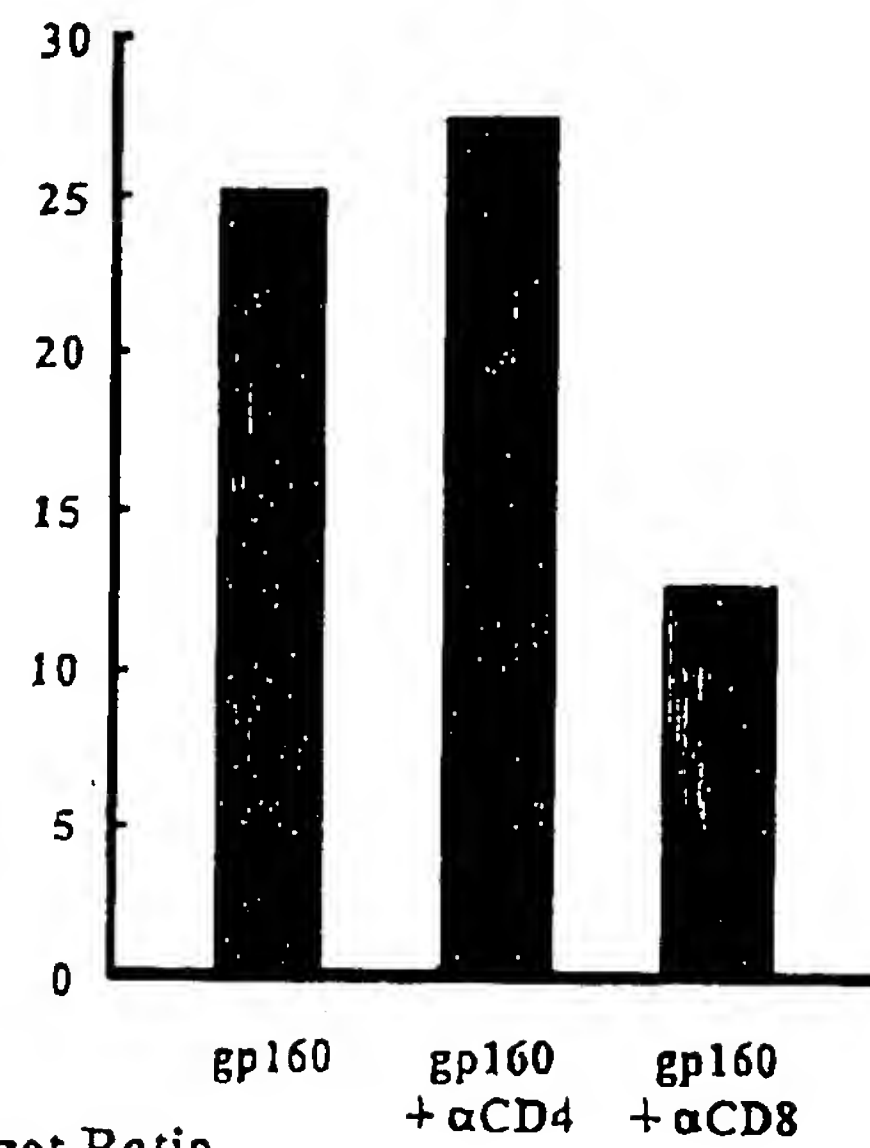
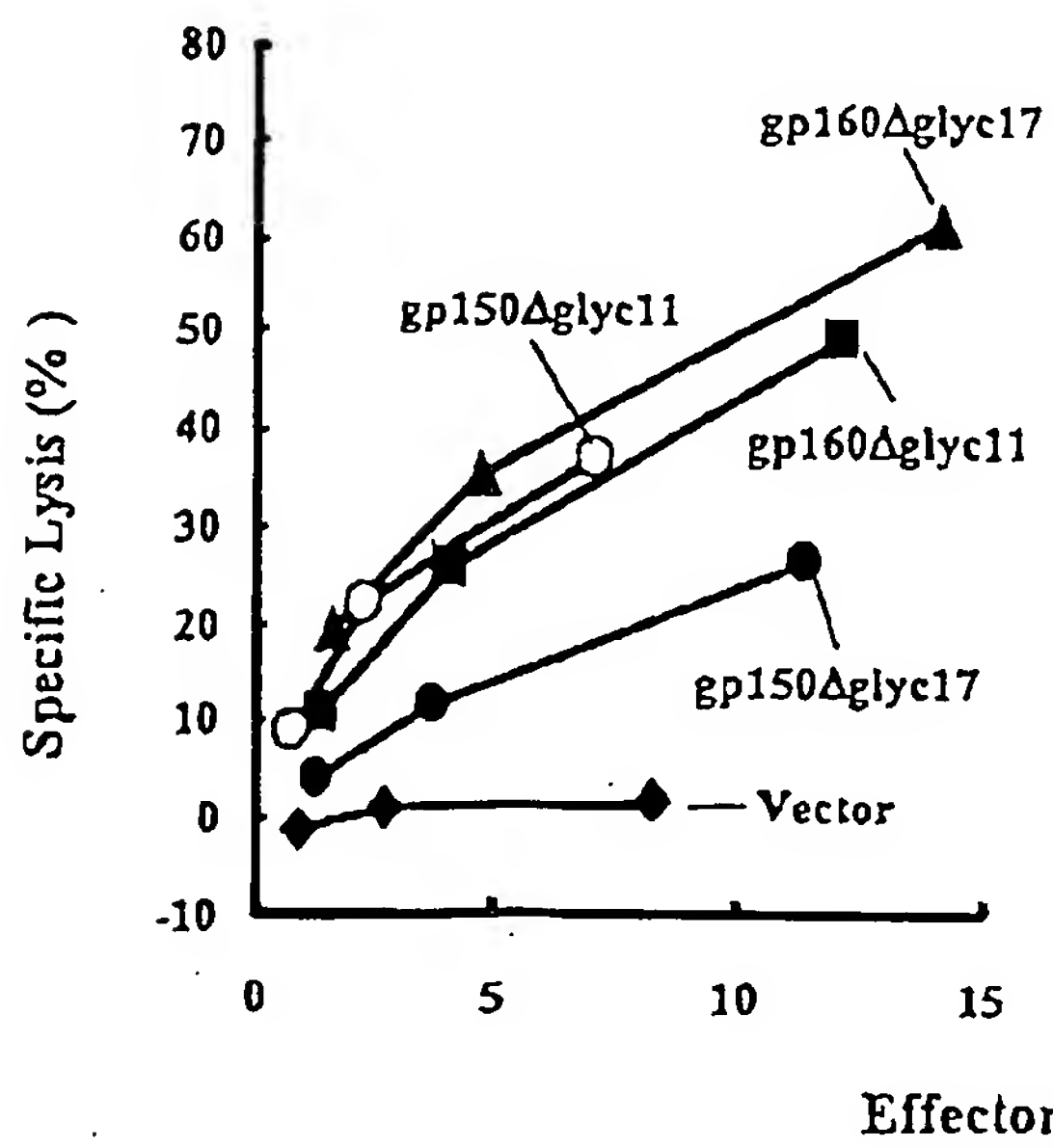
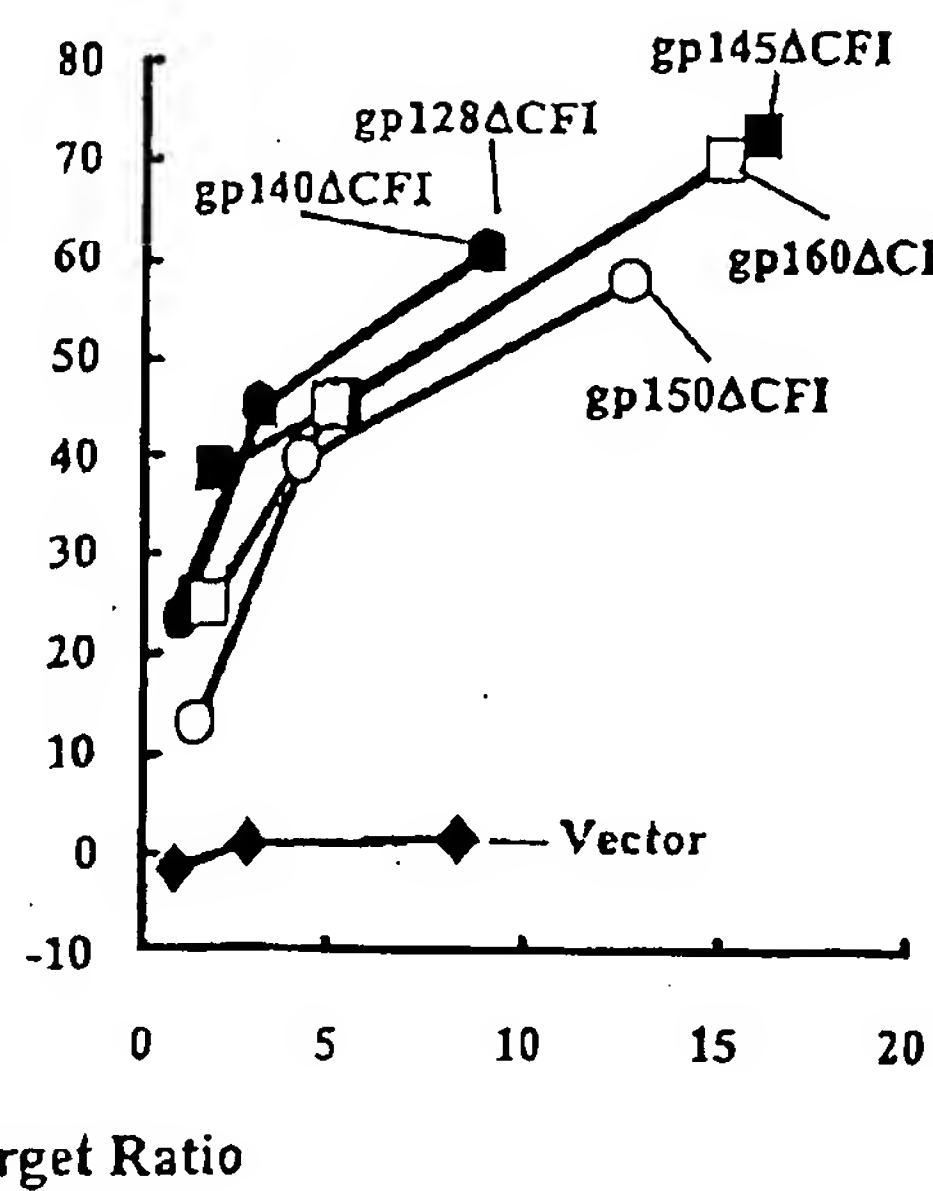
**A.****B.****C.****D.**

Figure VII

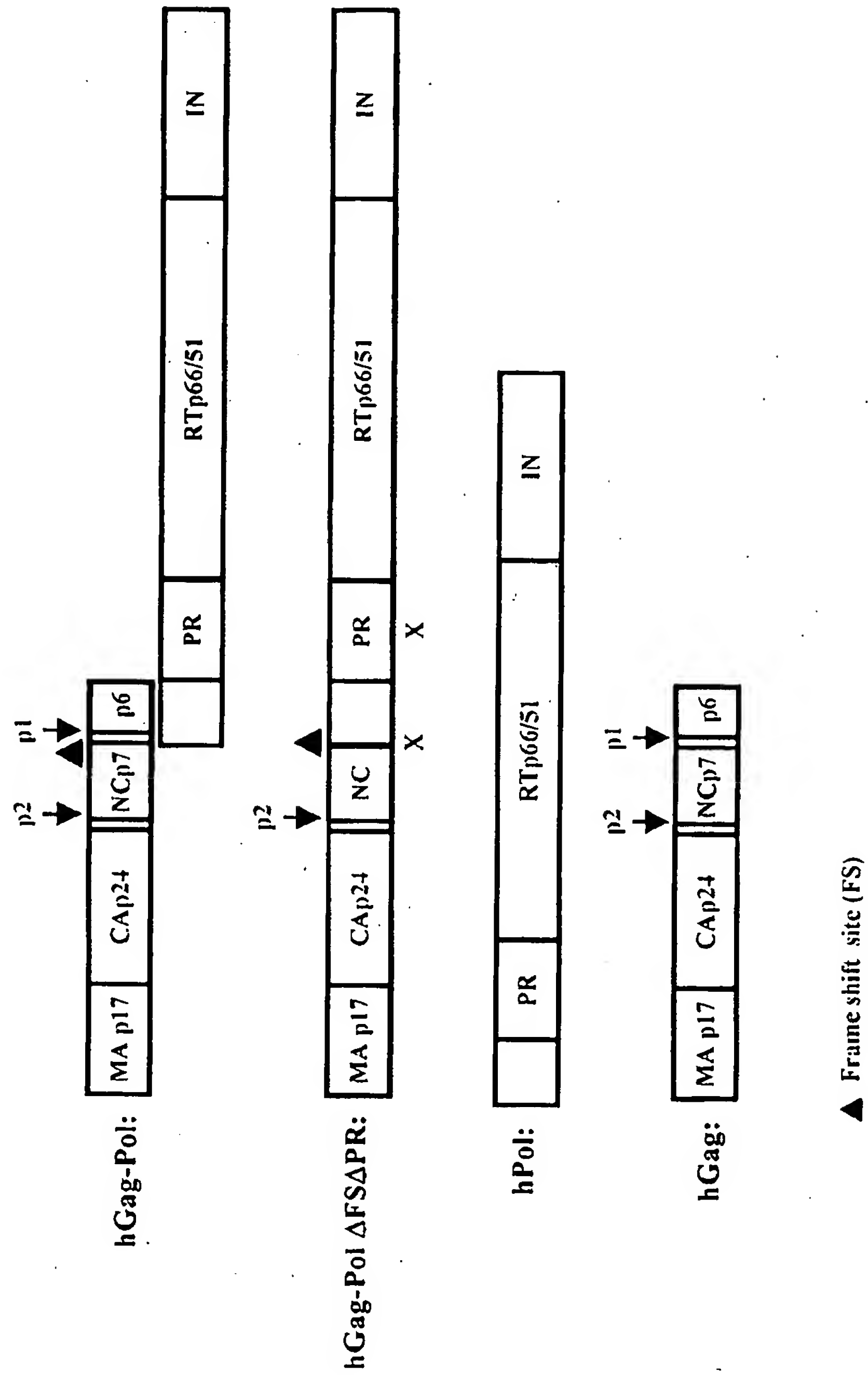
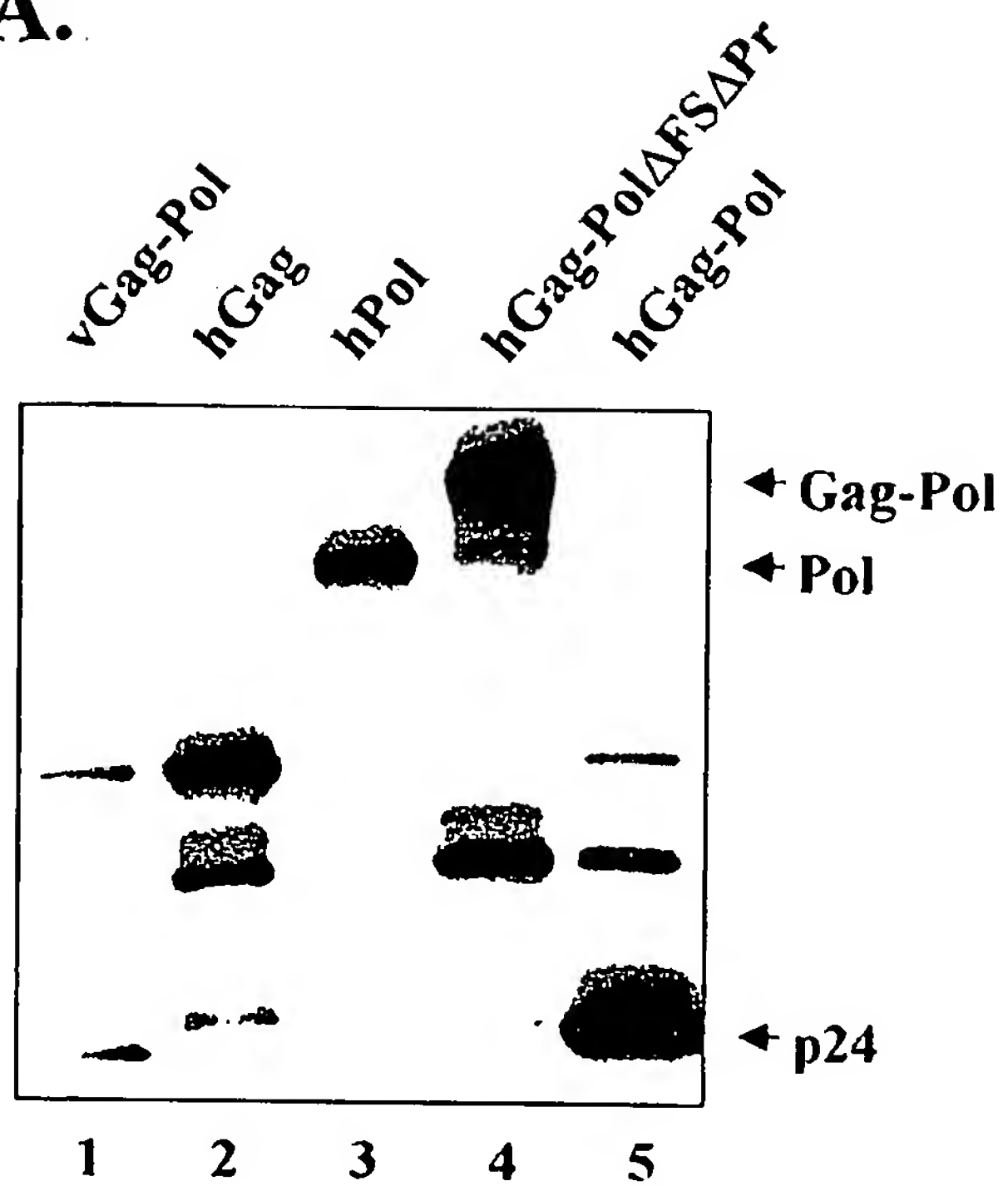
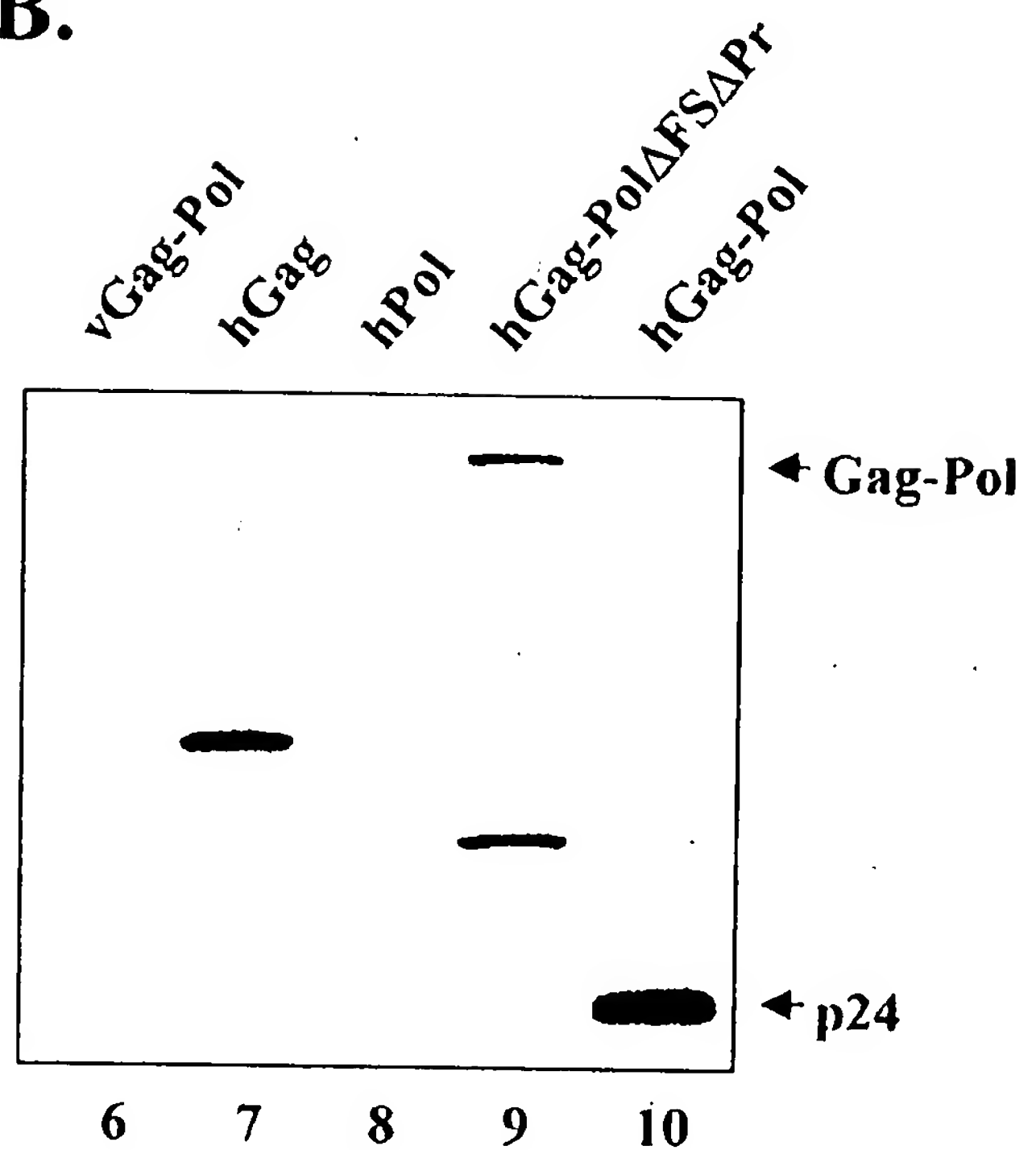


Figure VIII

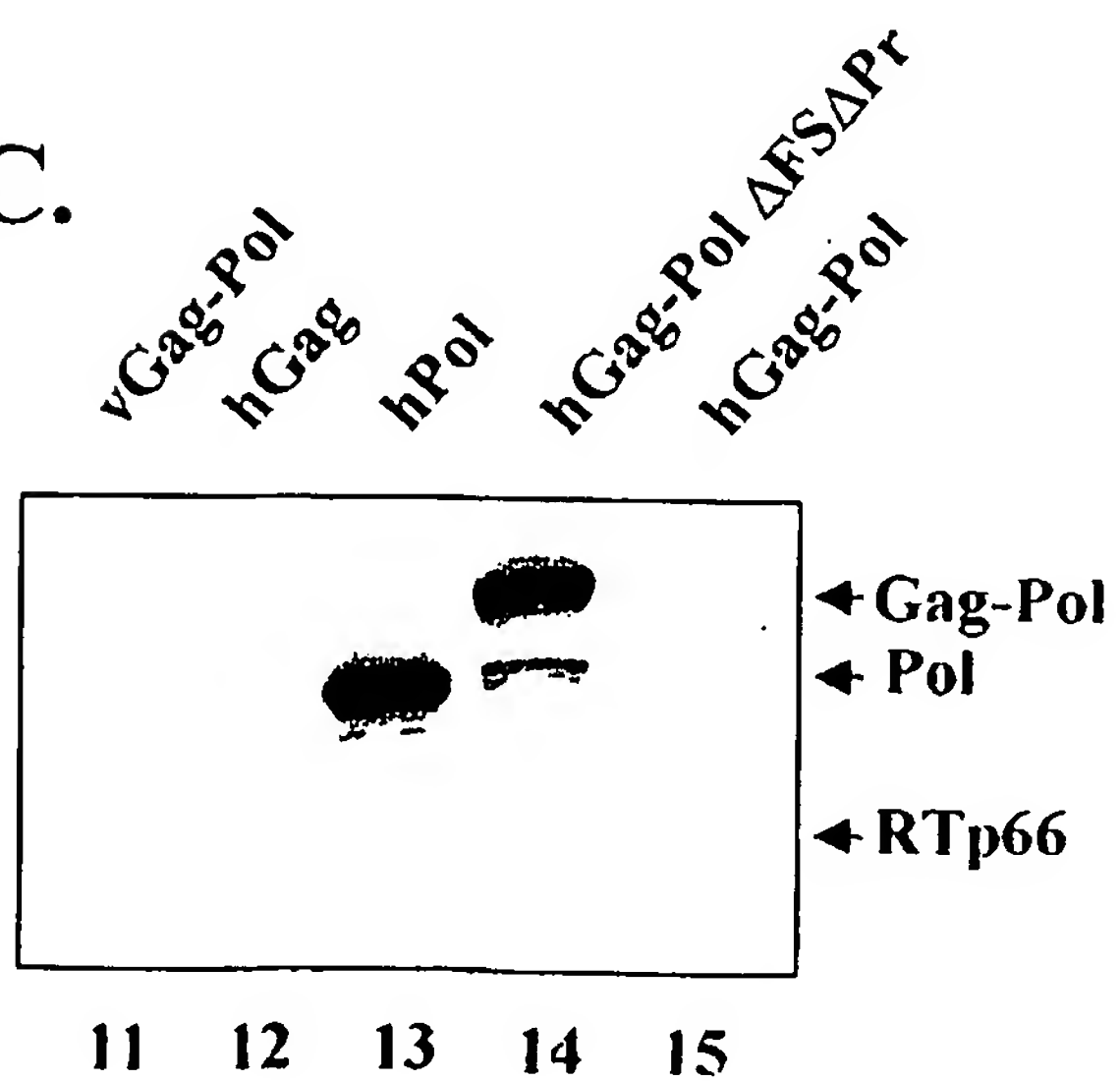
A.



B.



C.



D.

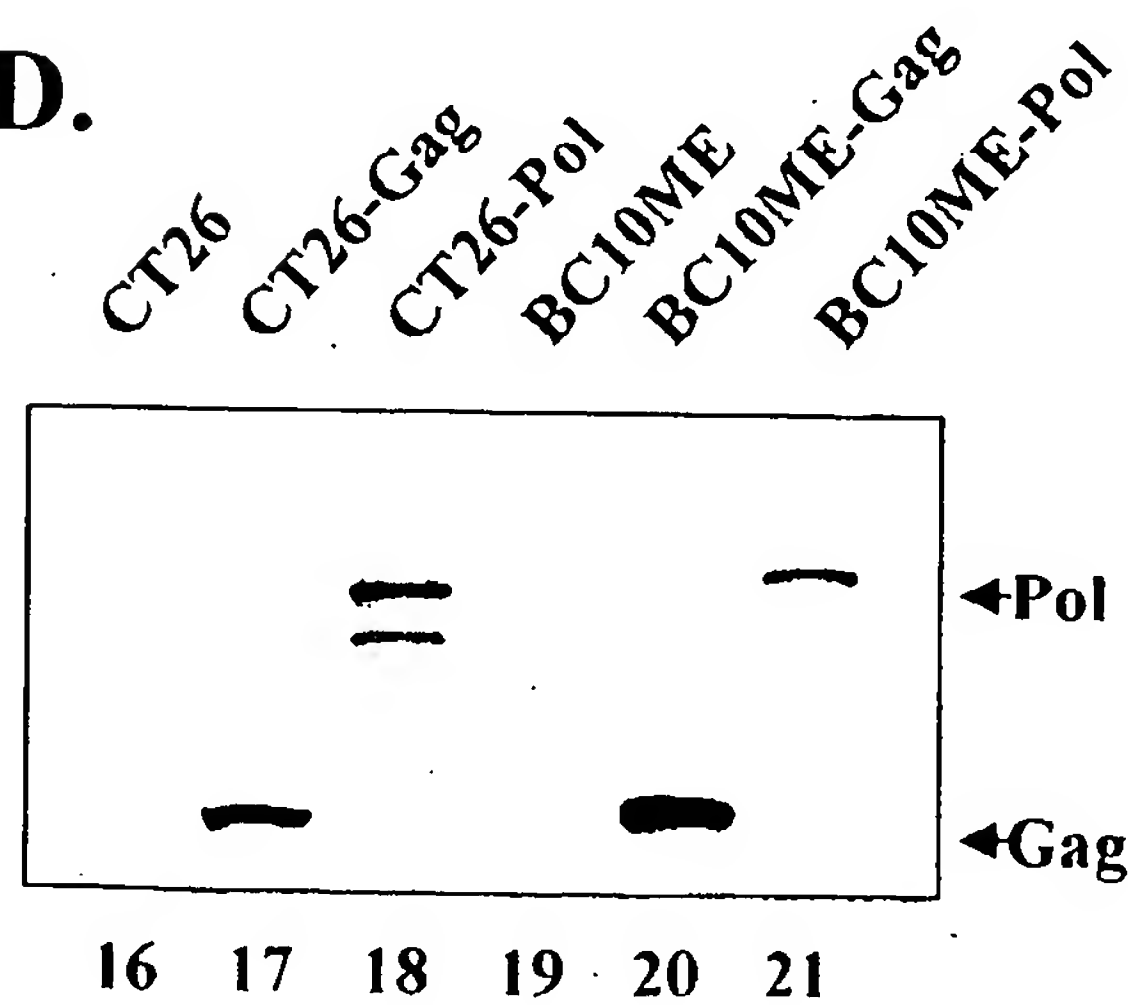
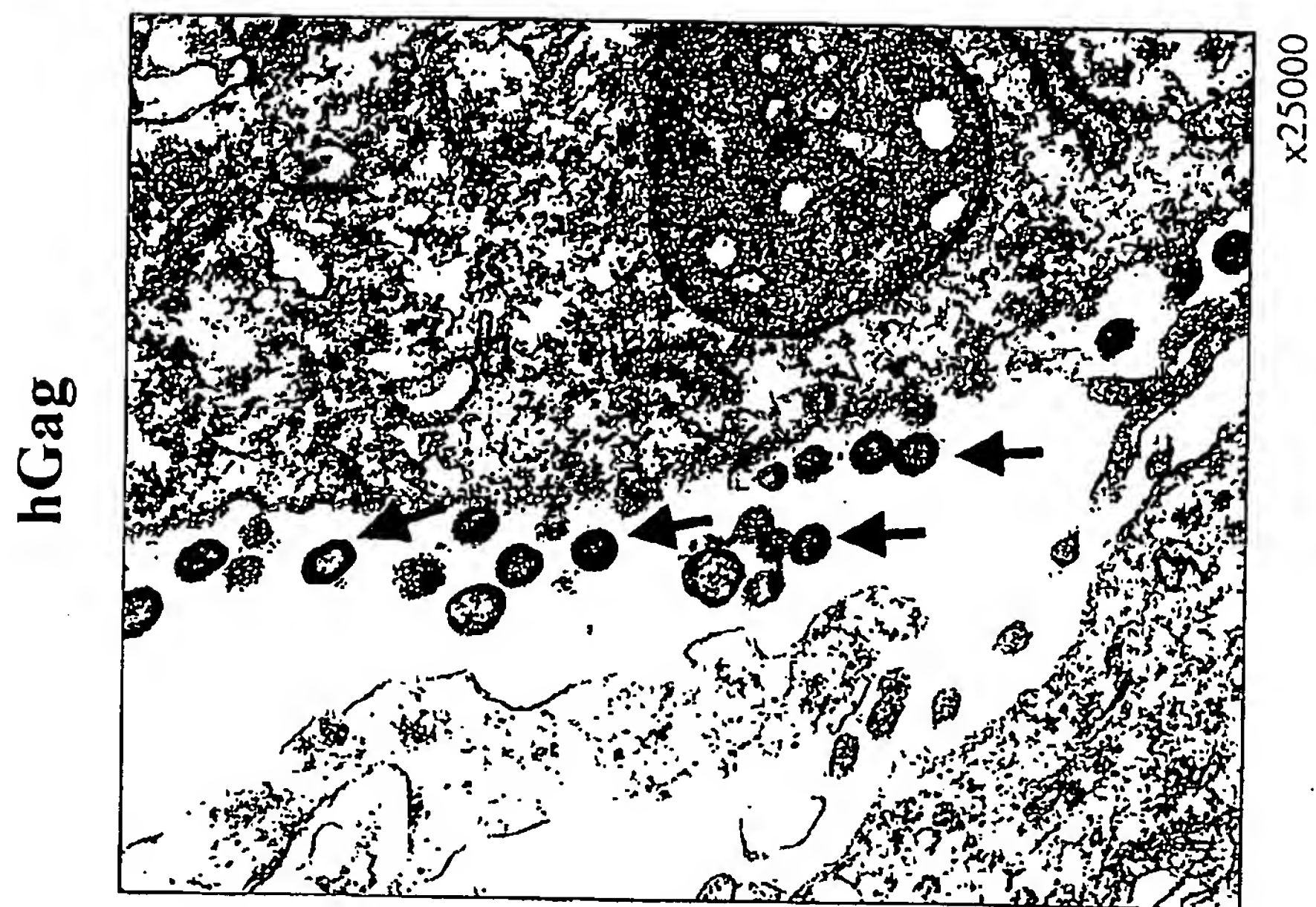
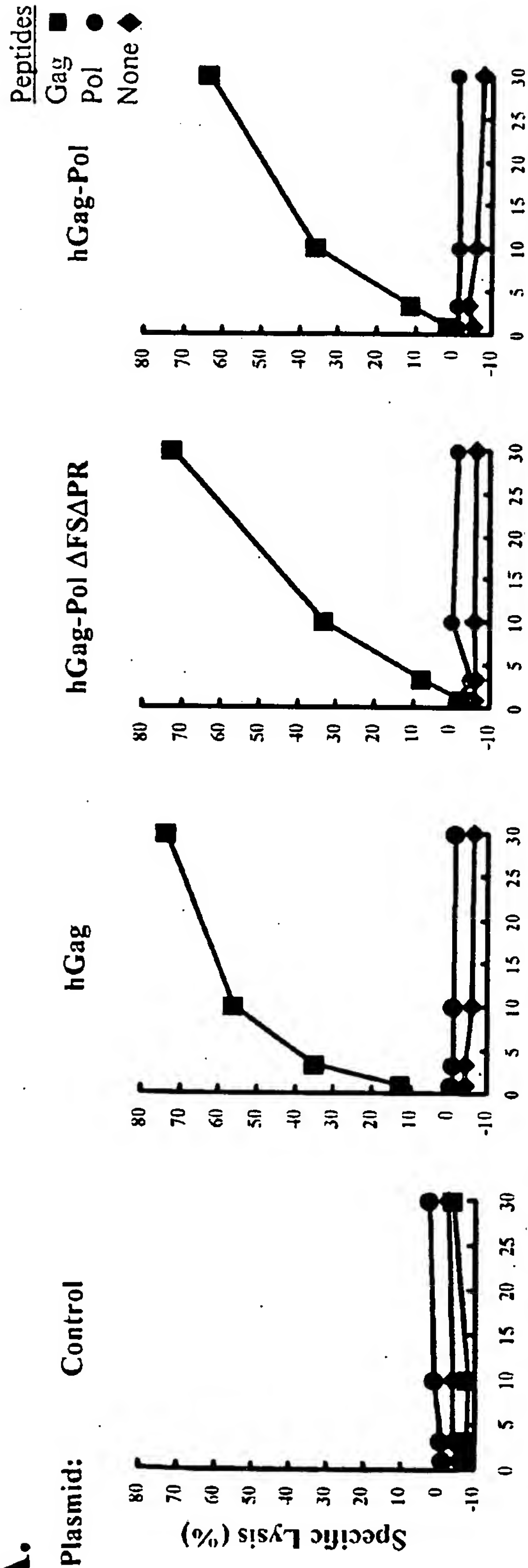


Figure IX



A.



B.

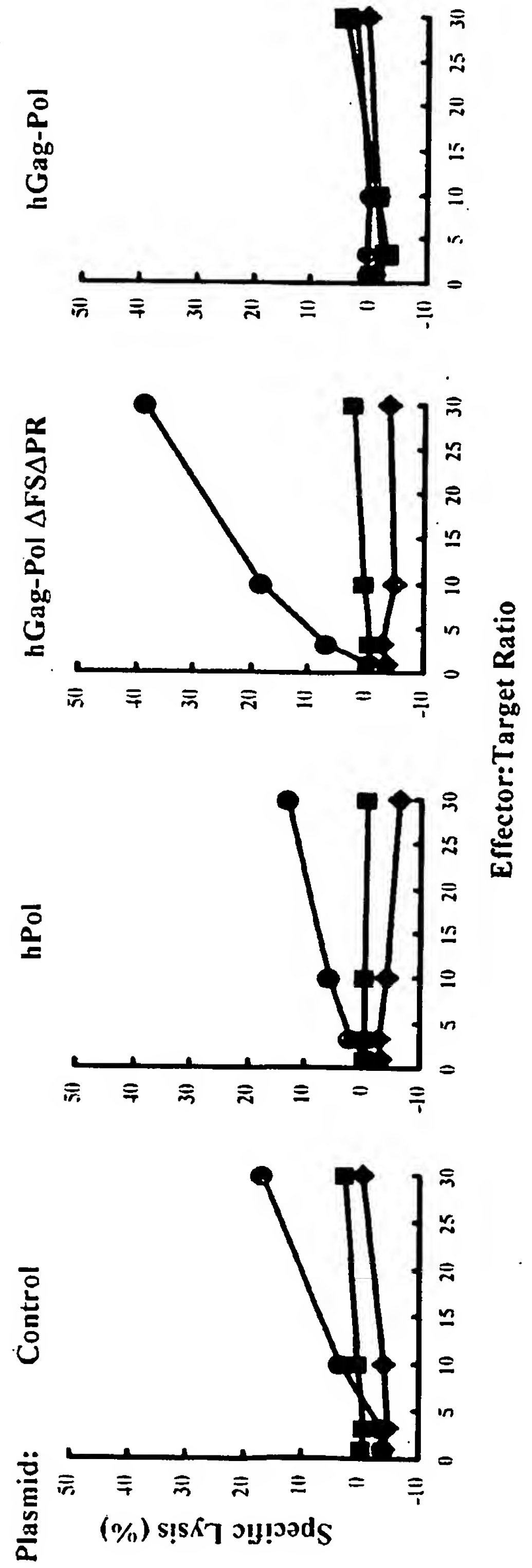




Figure X

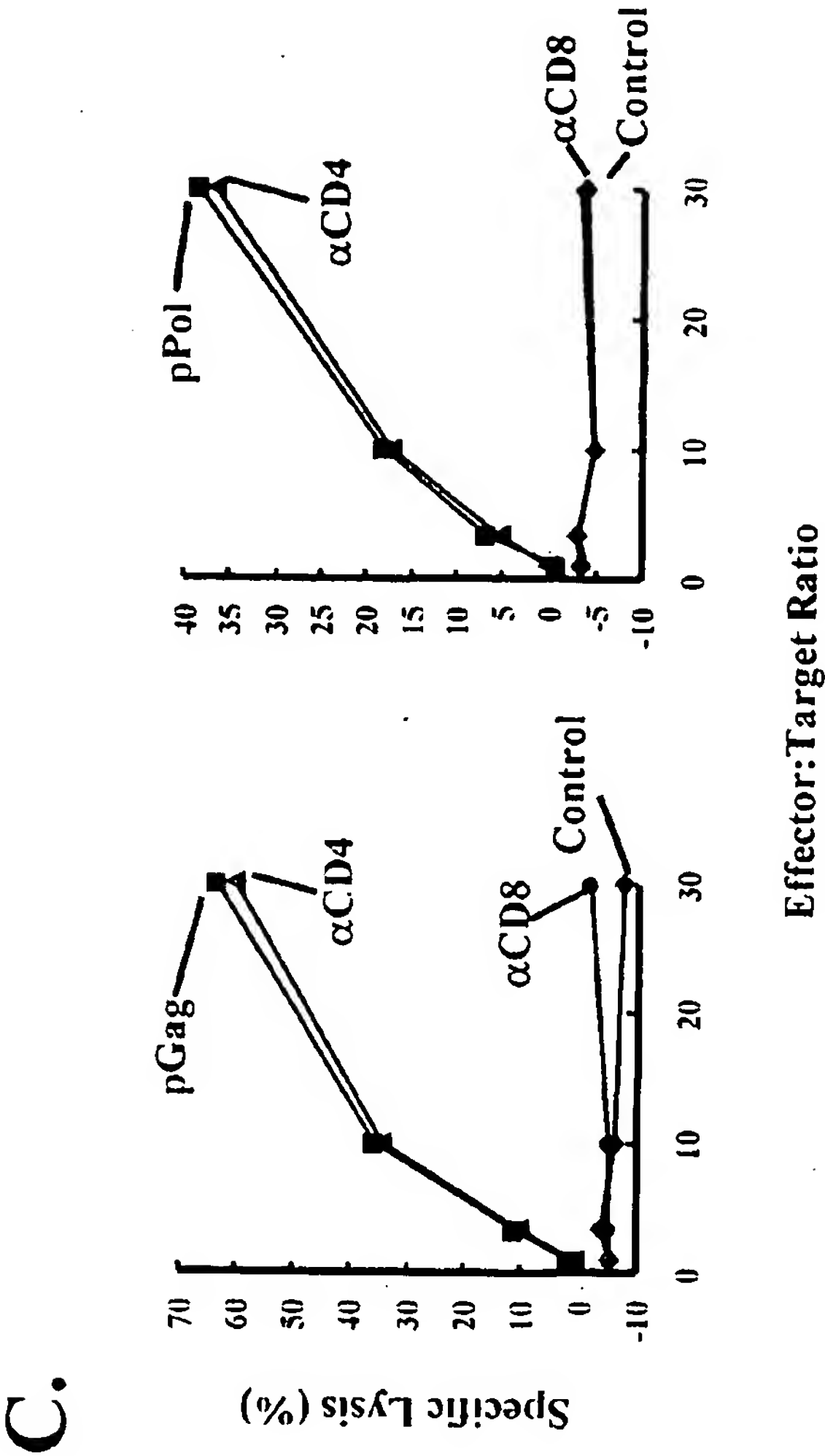
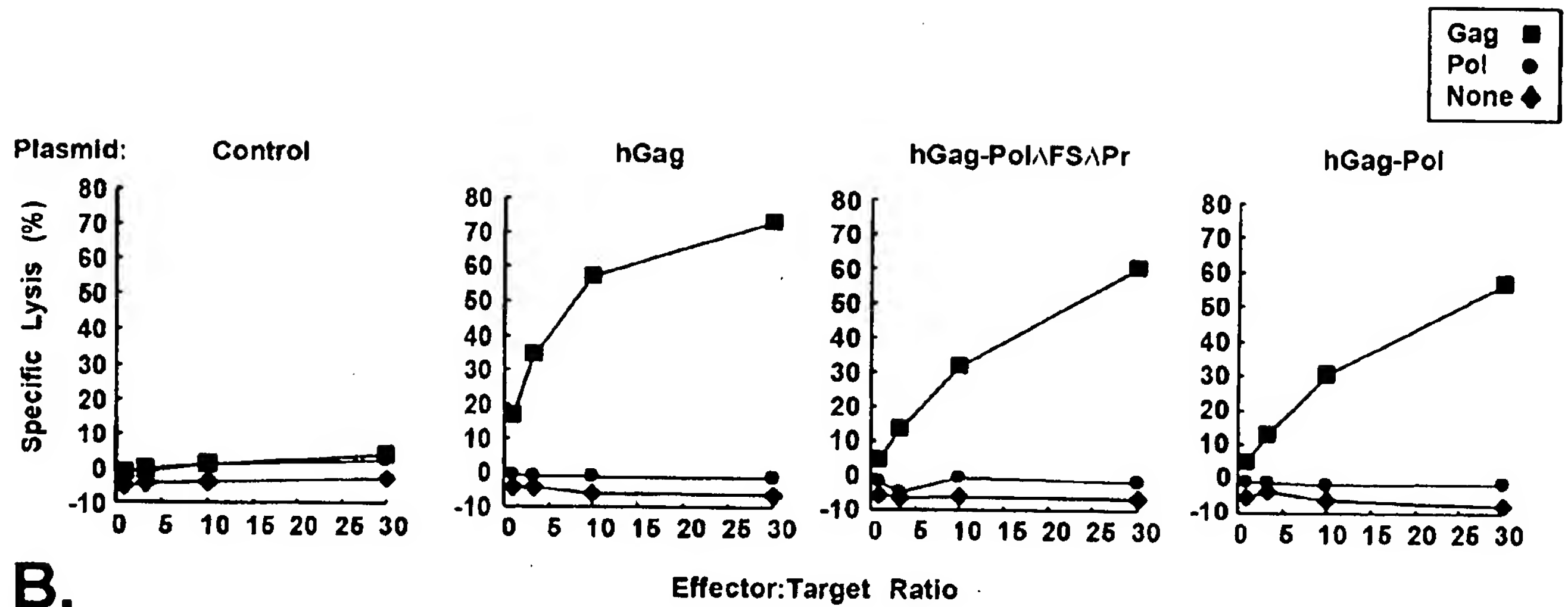
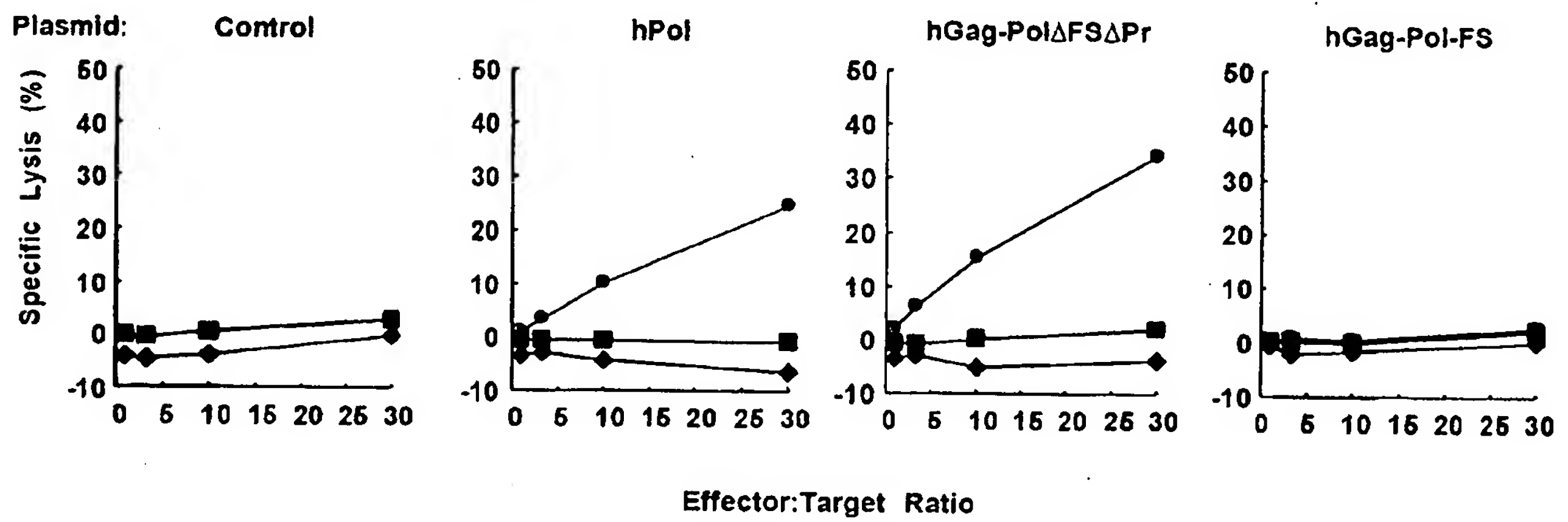
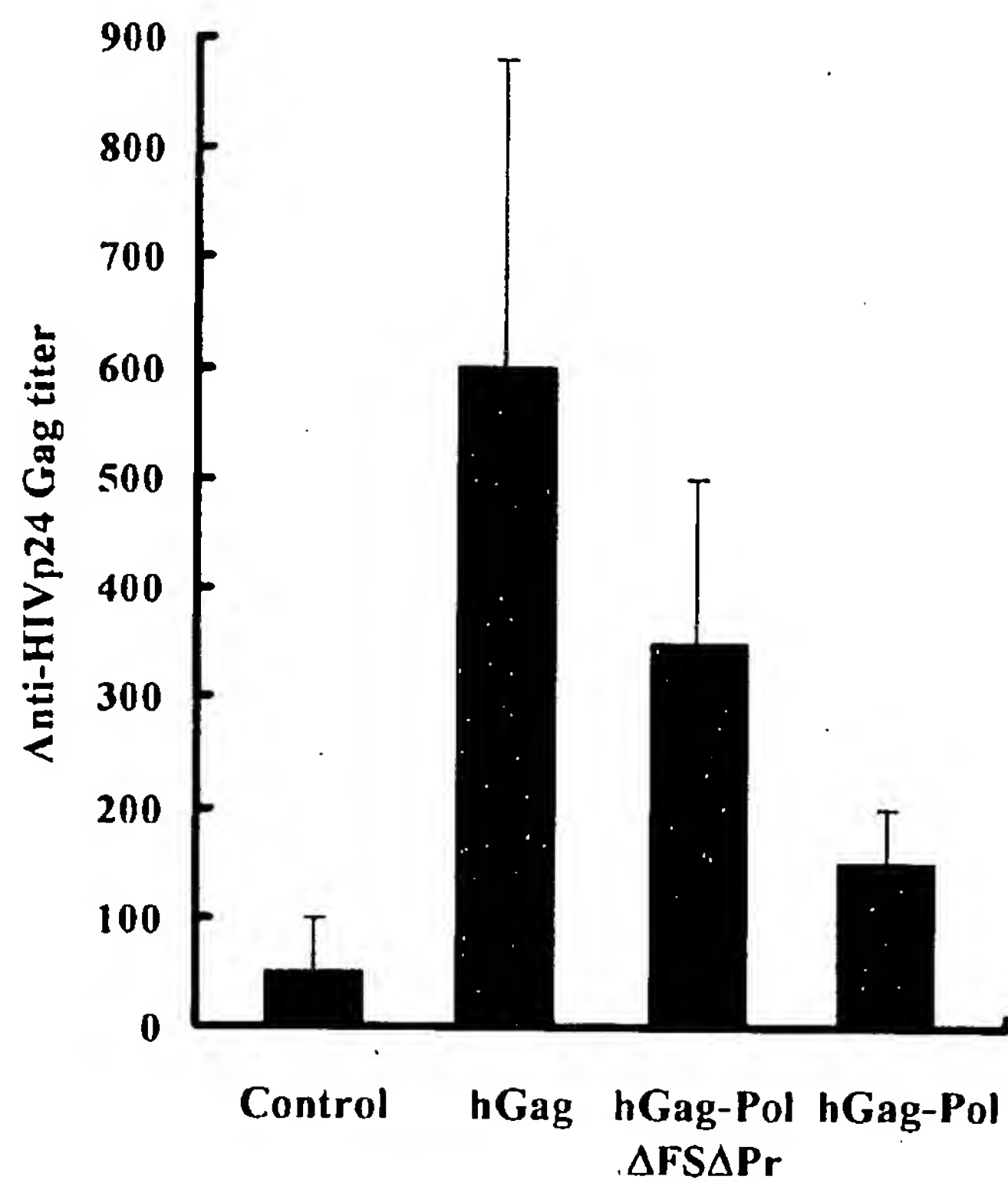
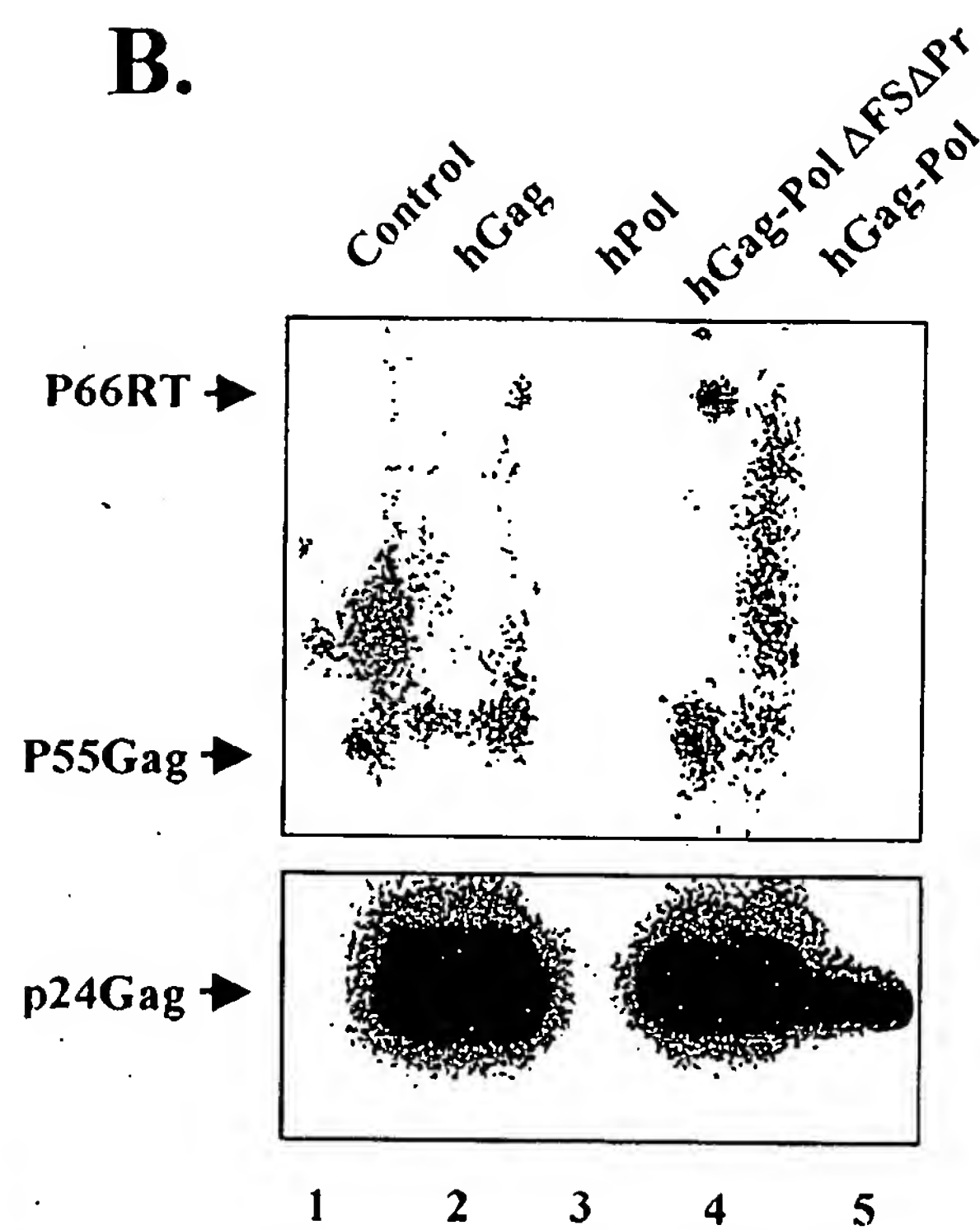
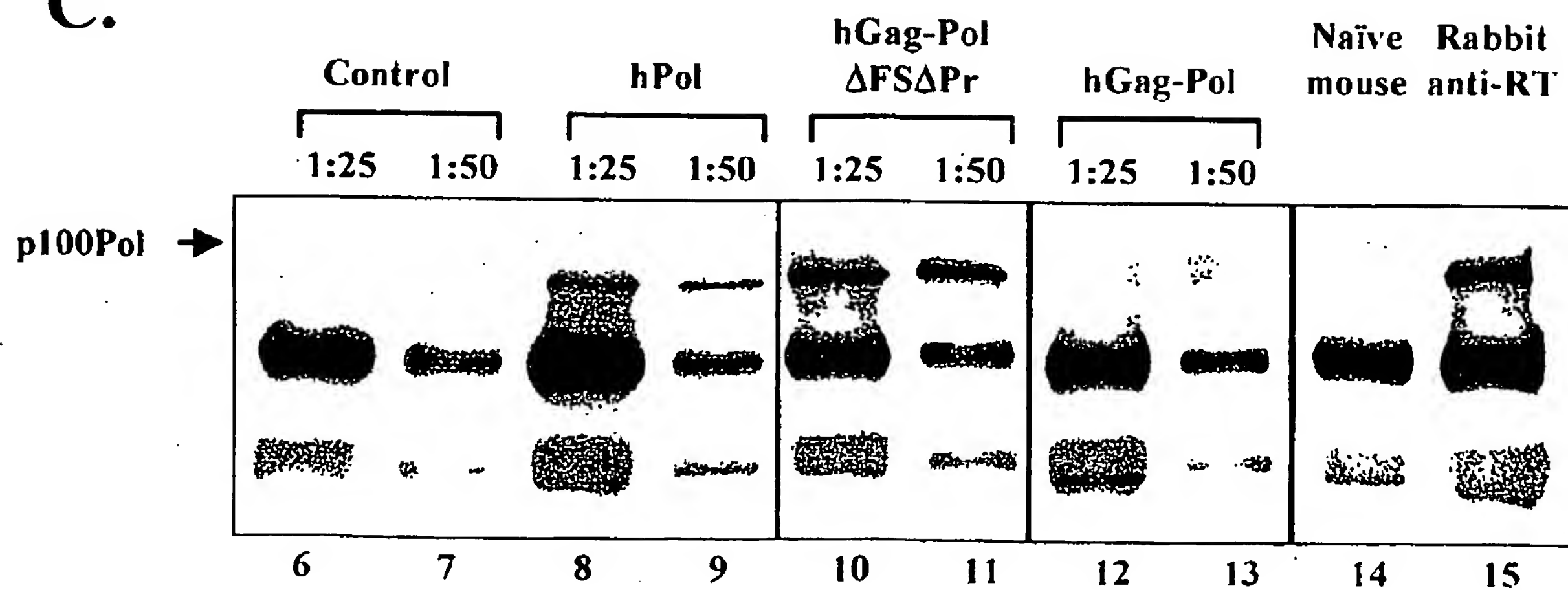


Figure XI

**A.****B.**

**A.****B.****C.**

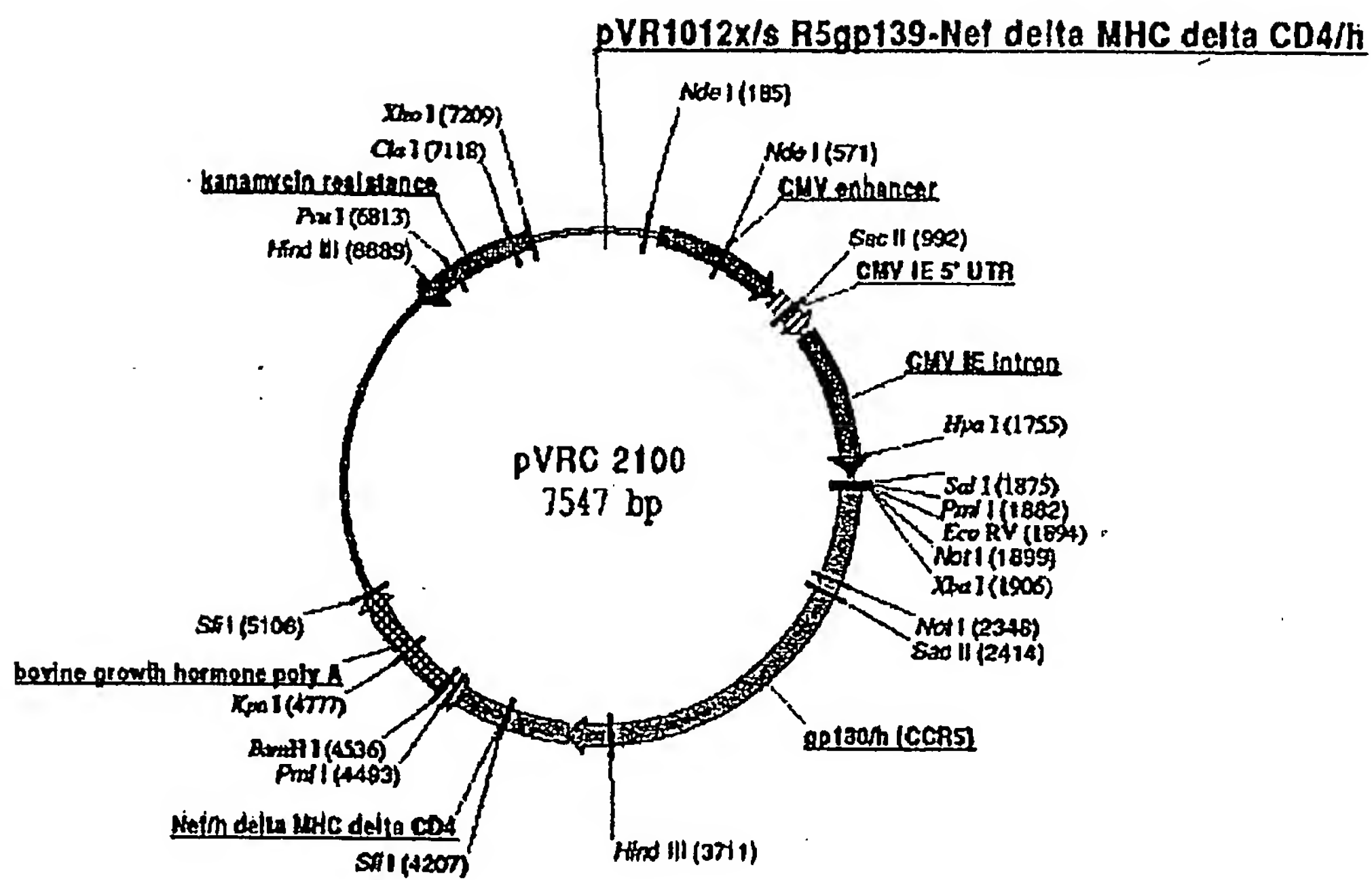


Figure 1

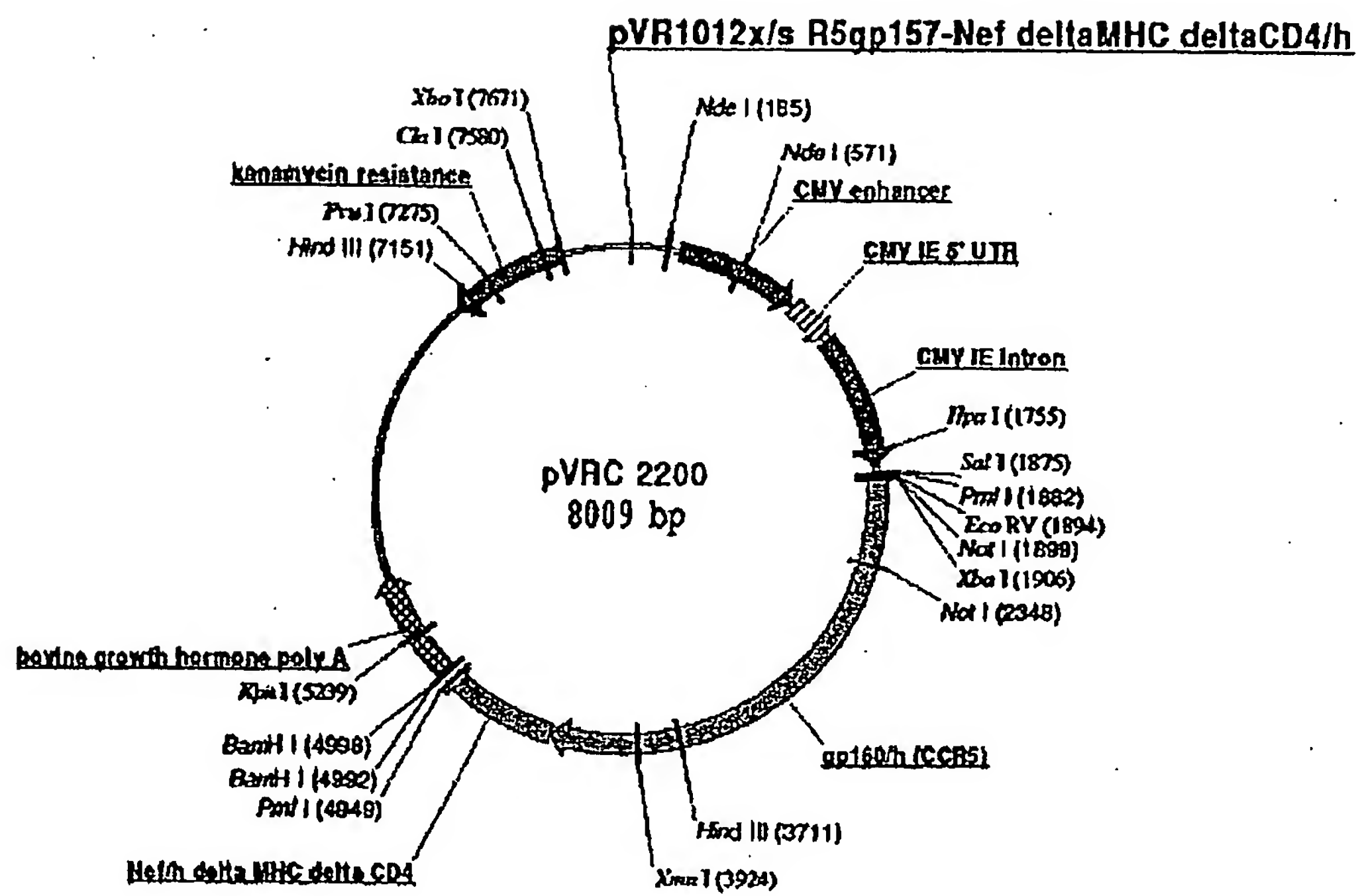


Figure 2



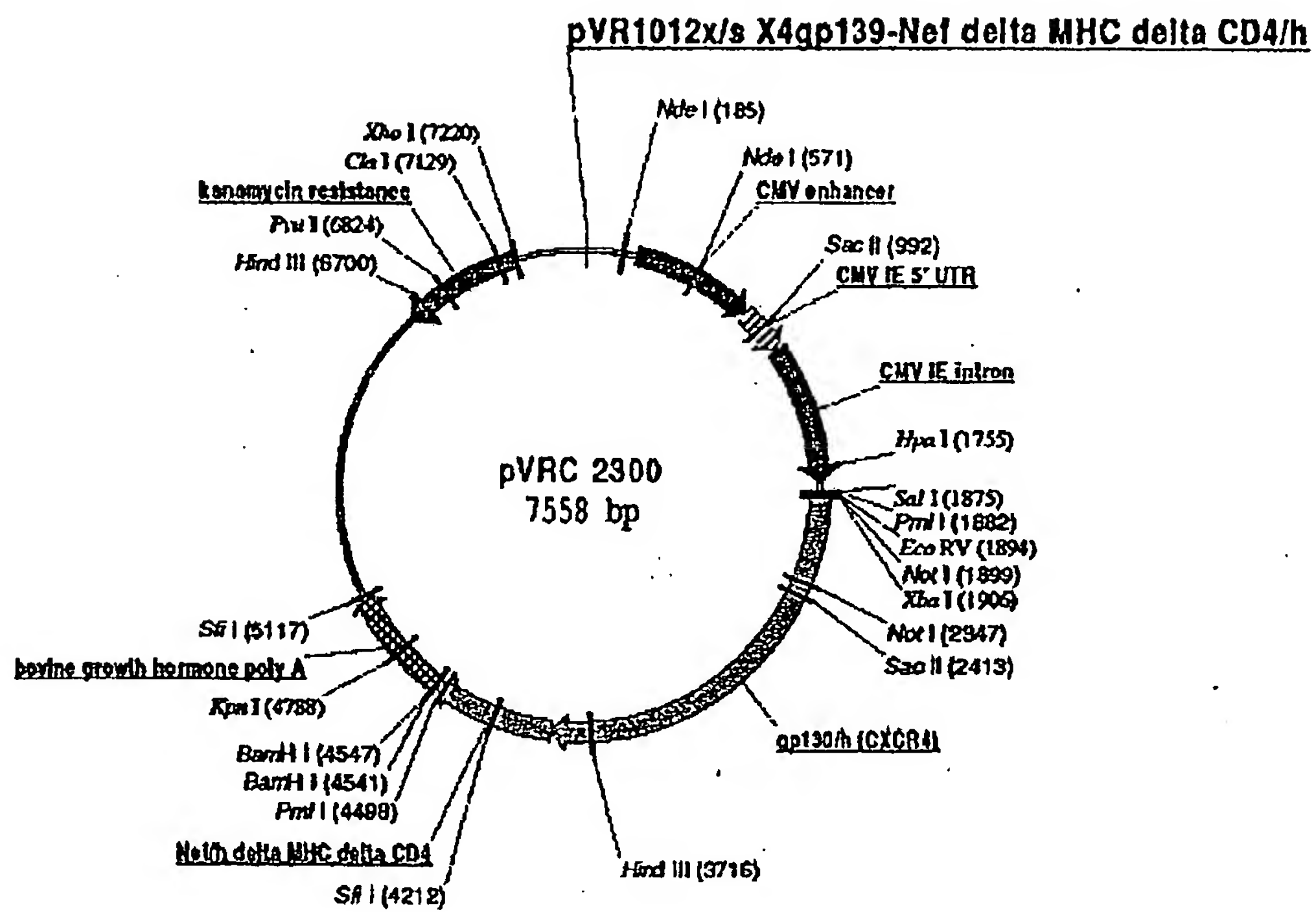


Figure 3

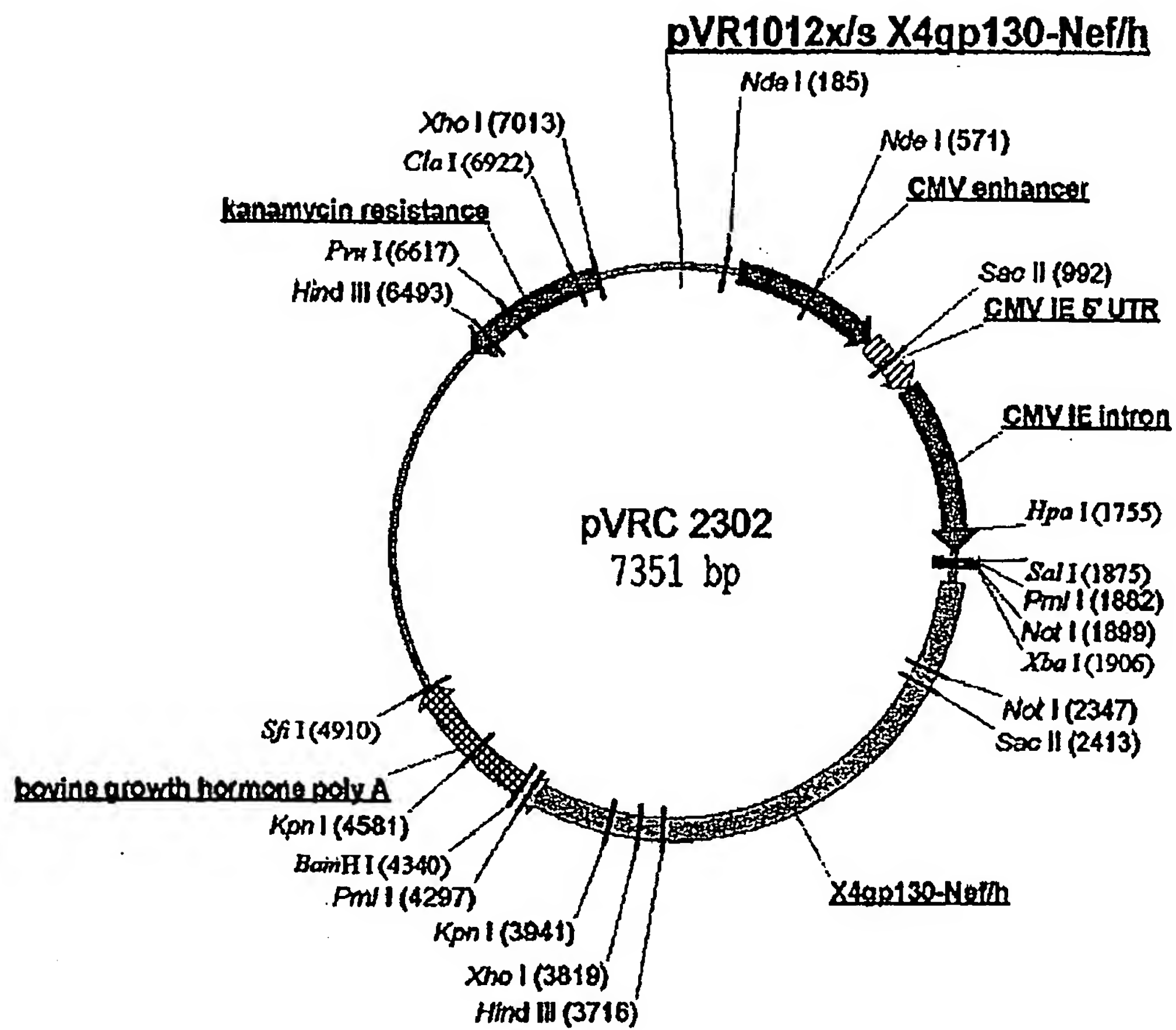


Figure 4

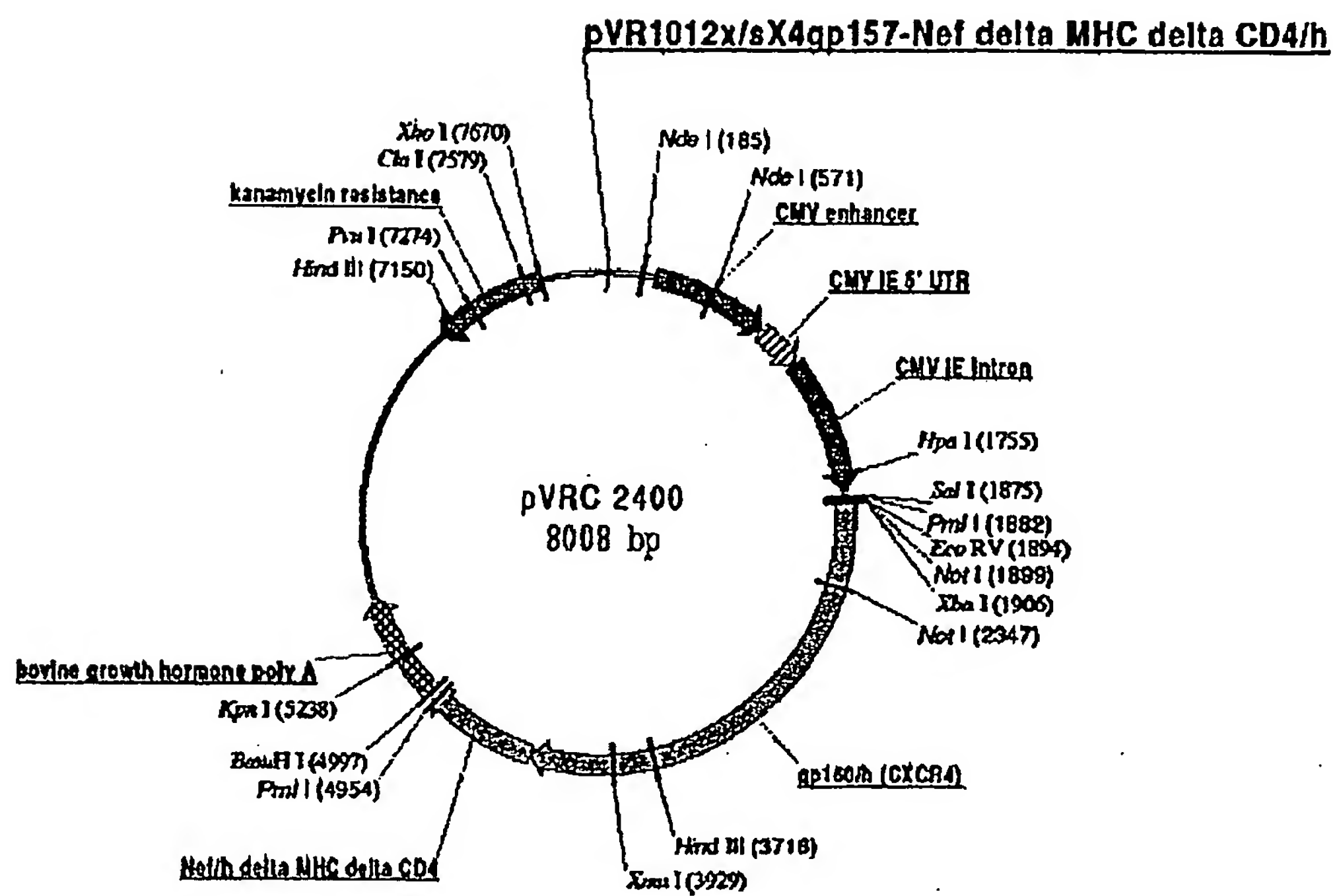


Figure 5

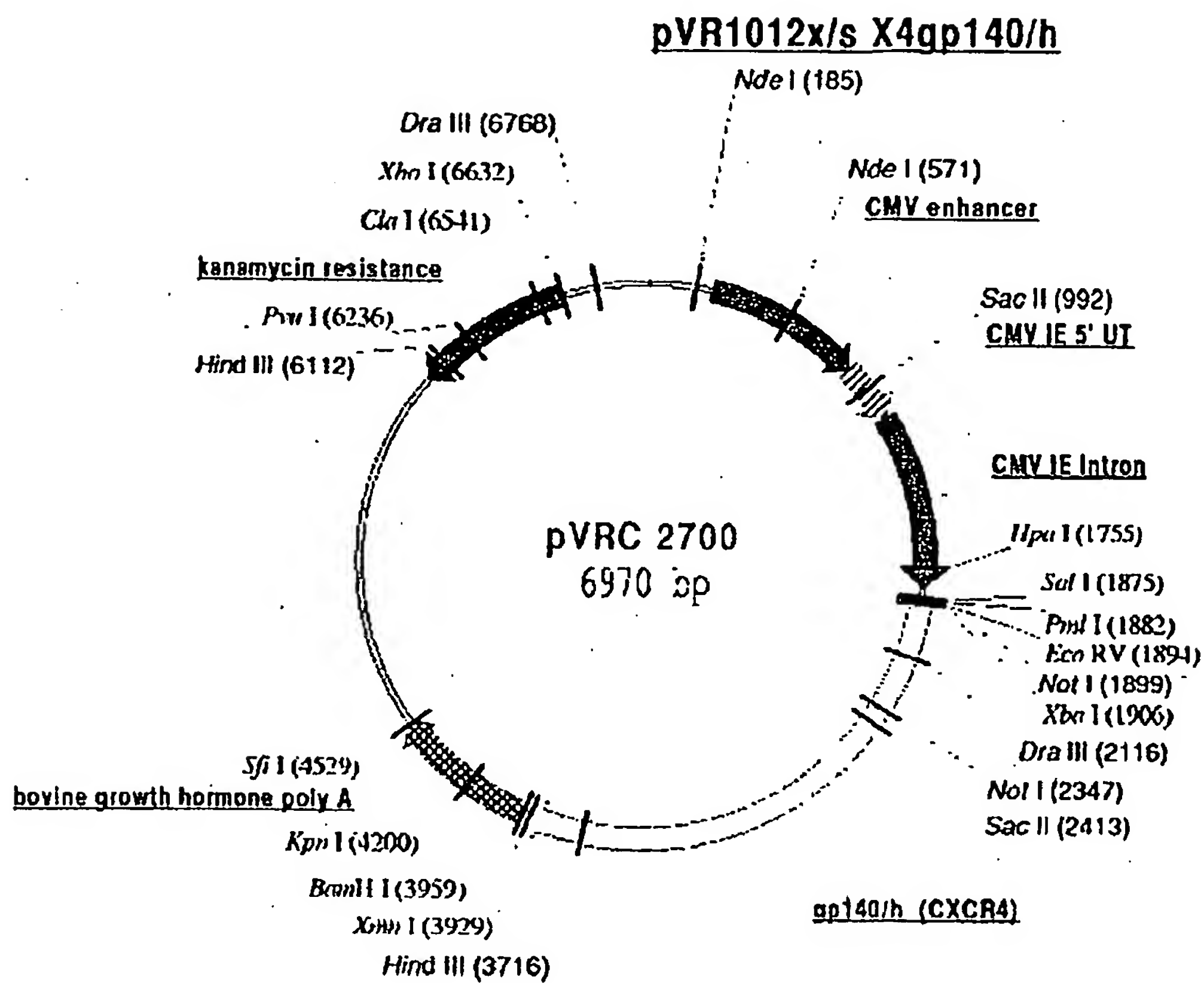


Figure 6

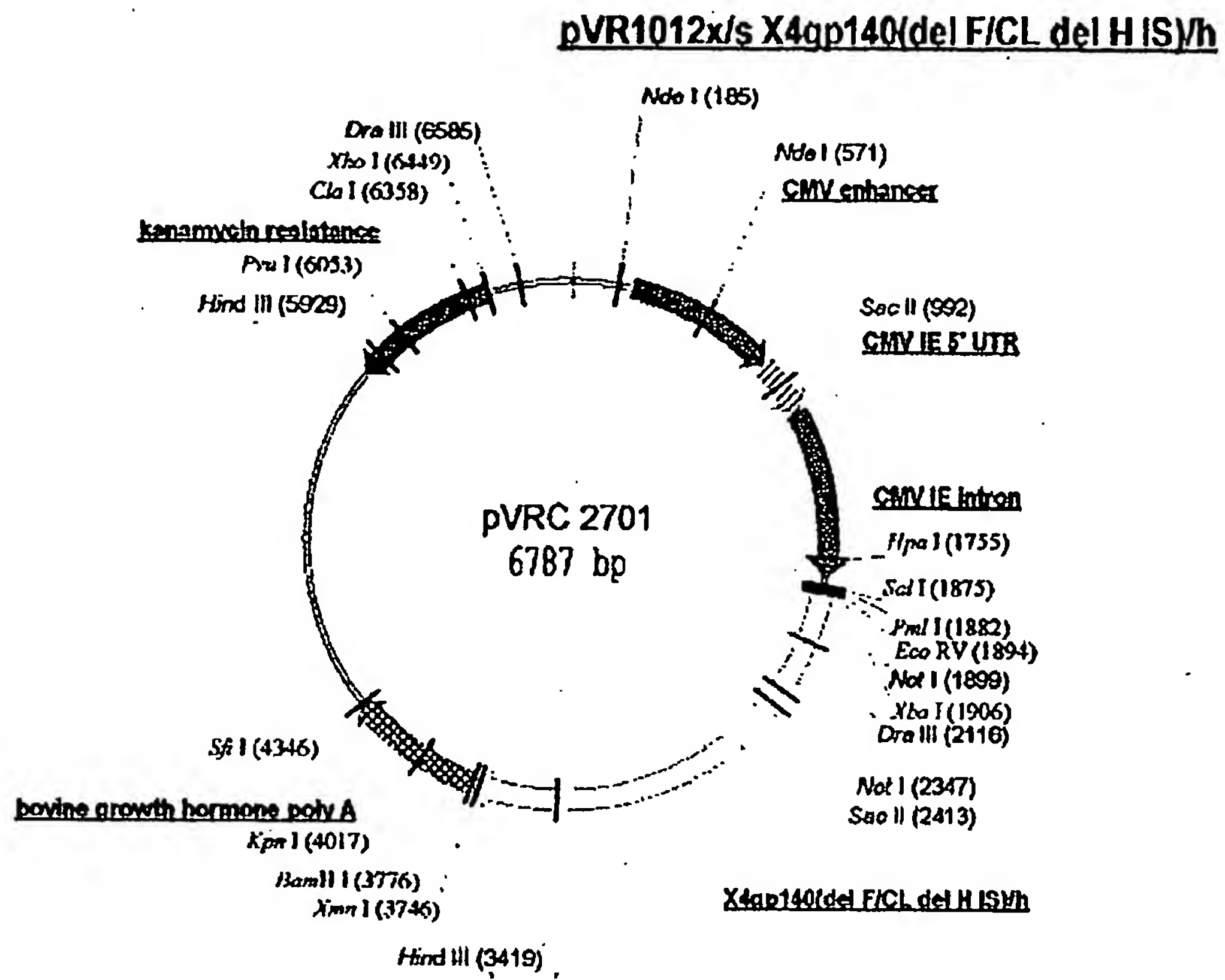


Figure 7



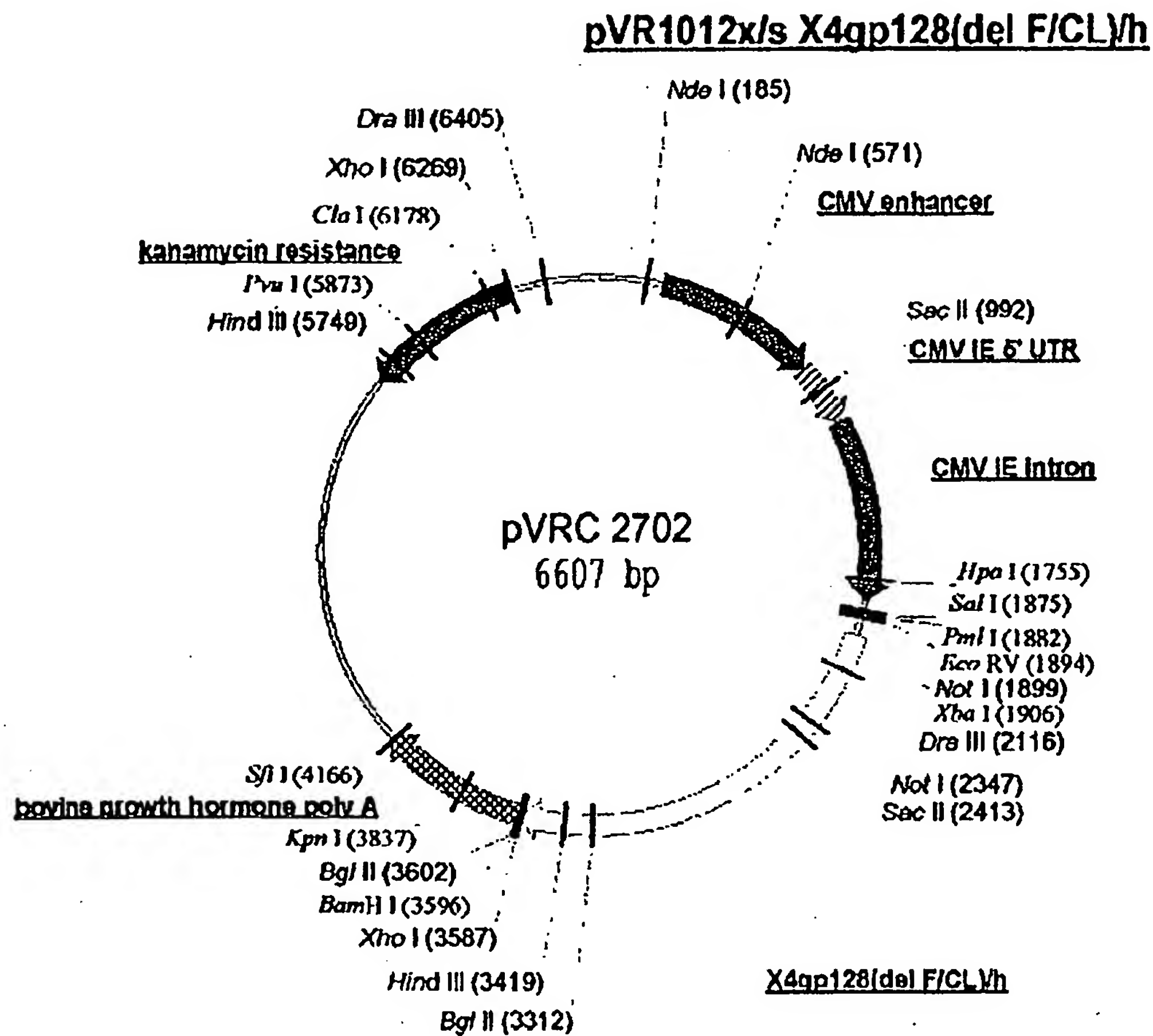


Figure 8

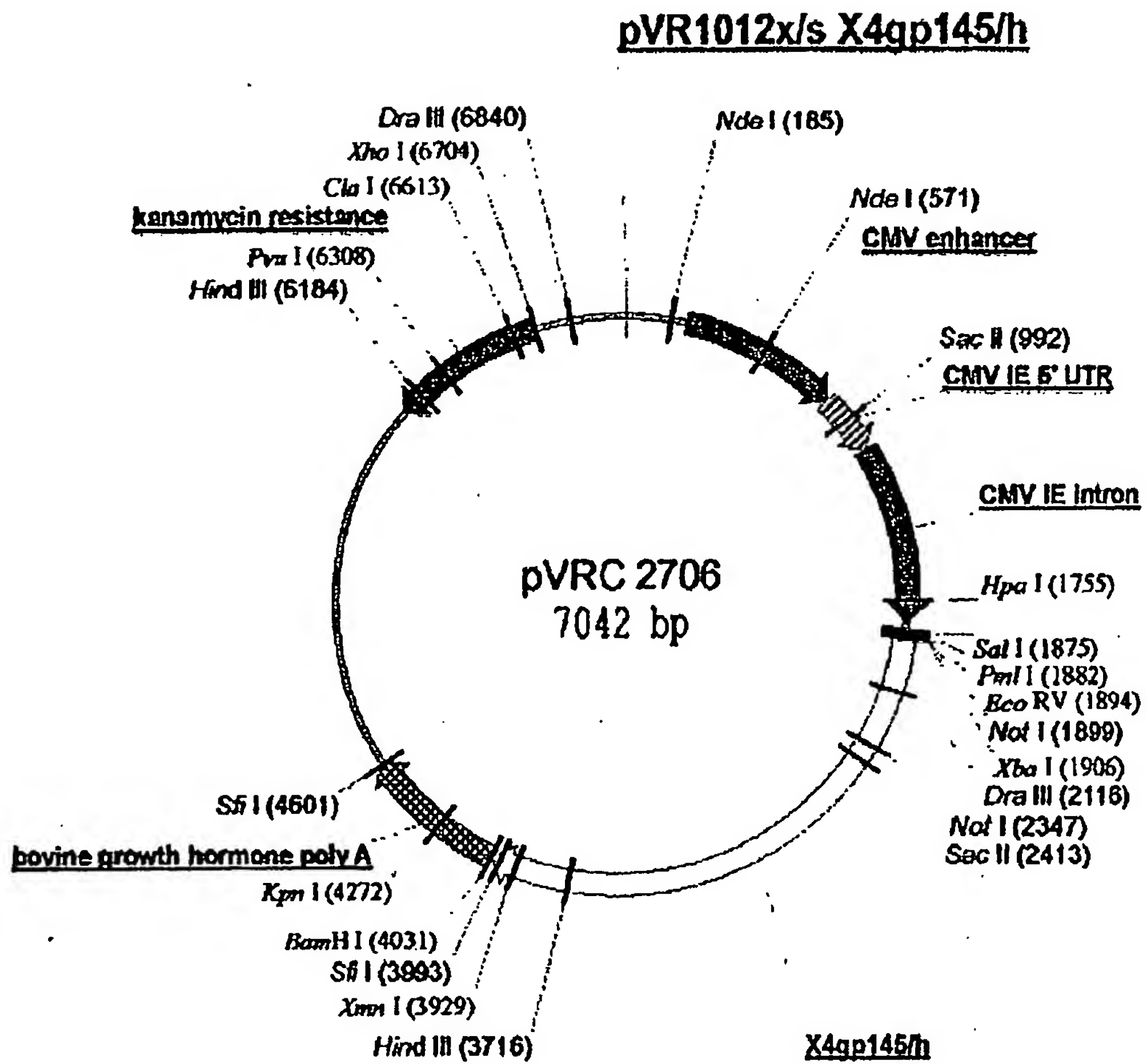


Figure 9

**pVR1012x/s X4gp145(del F/CL del H IS)/h**

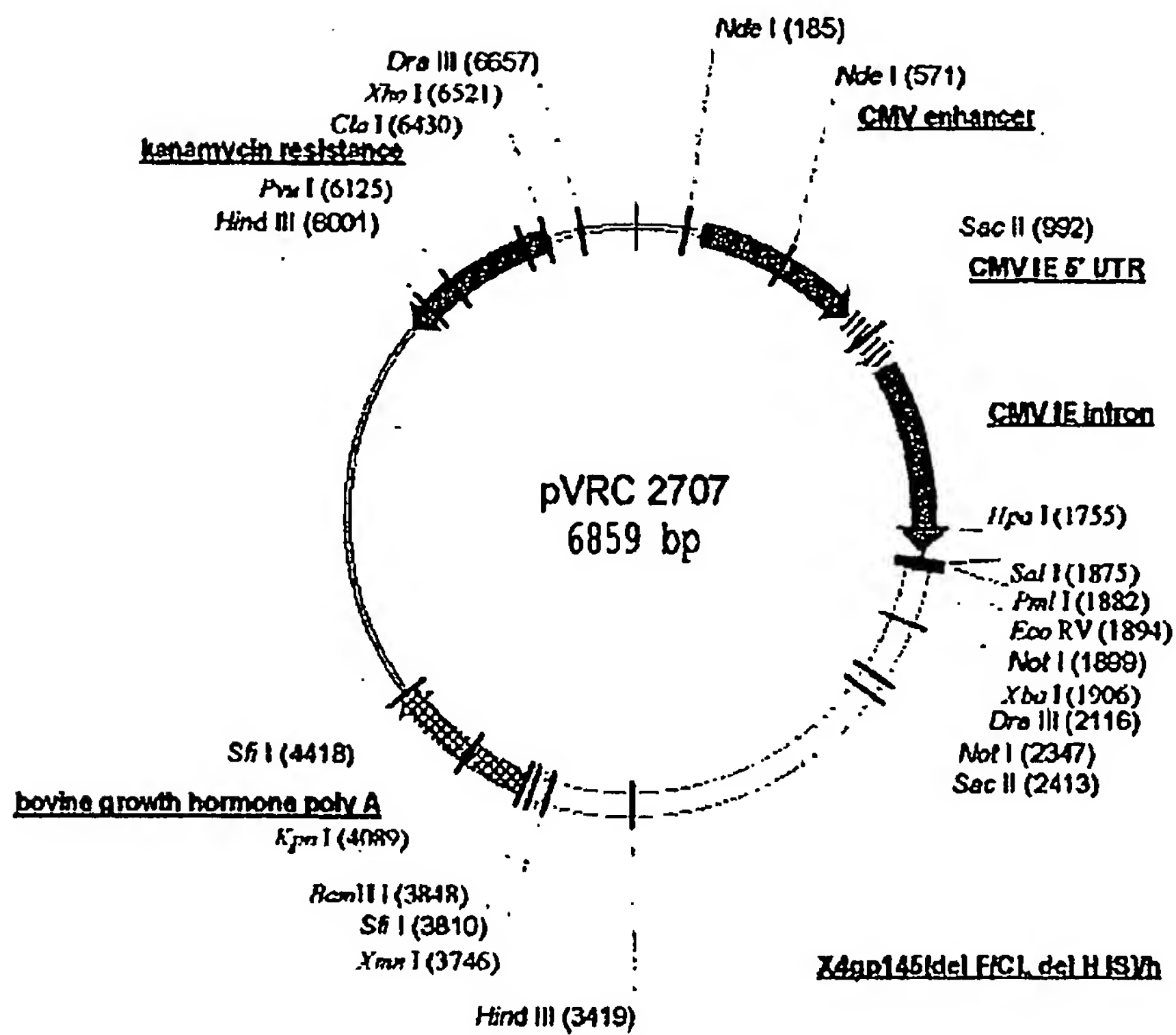


Figure 10

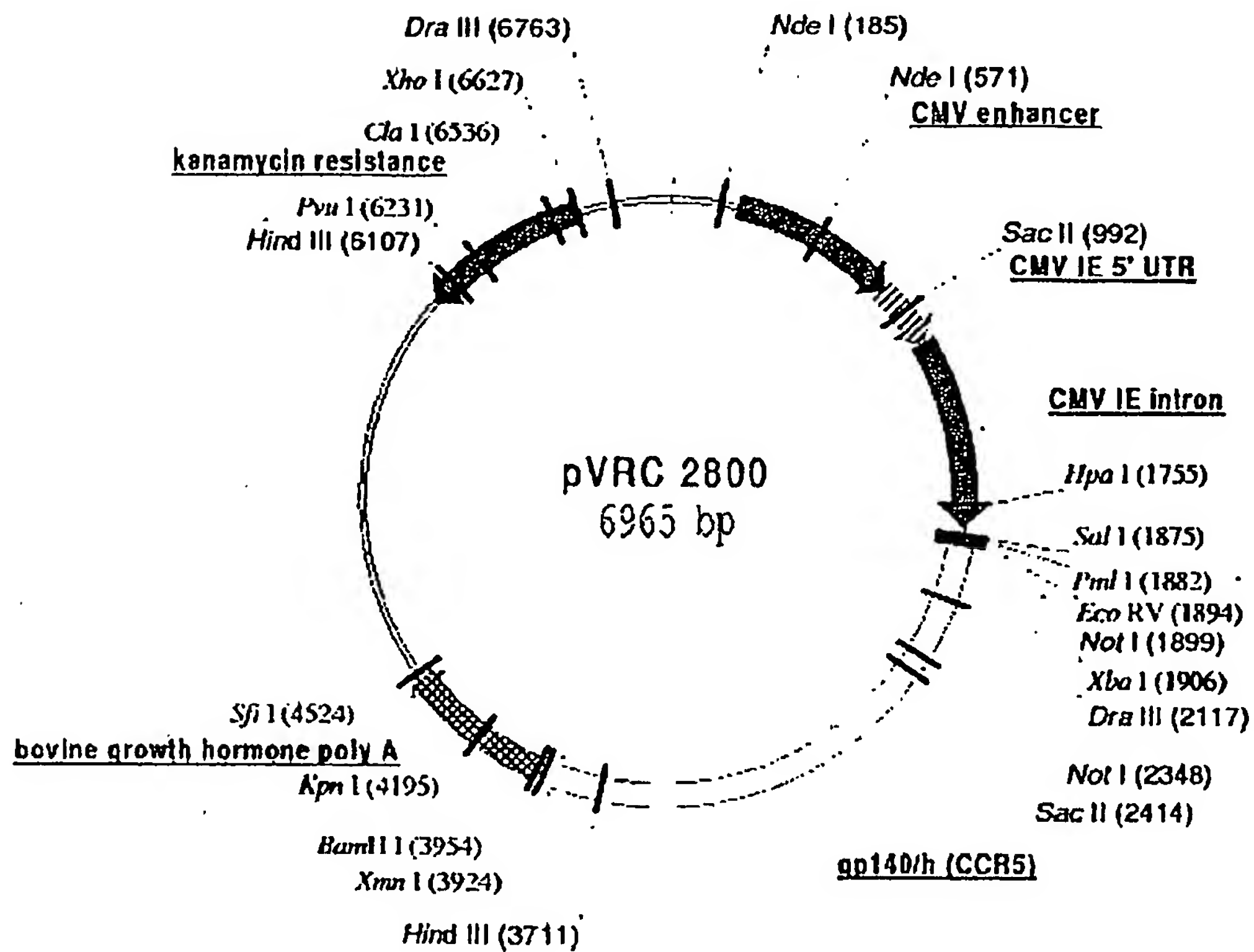
**pVR1012x/s R5gp140/h**

Figure 11

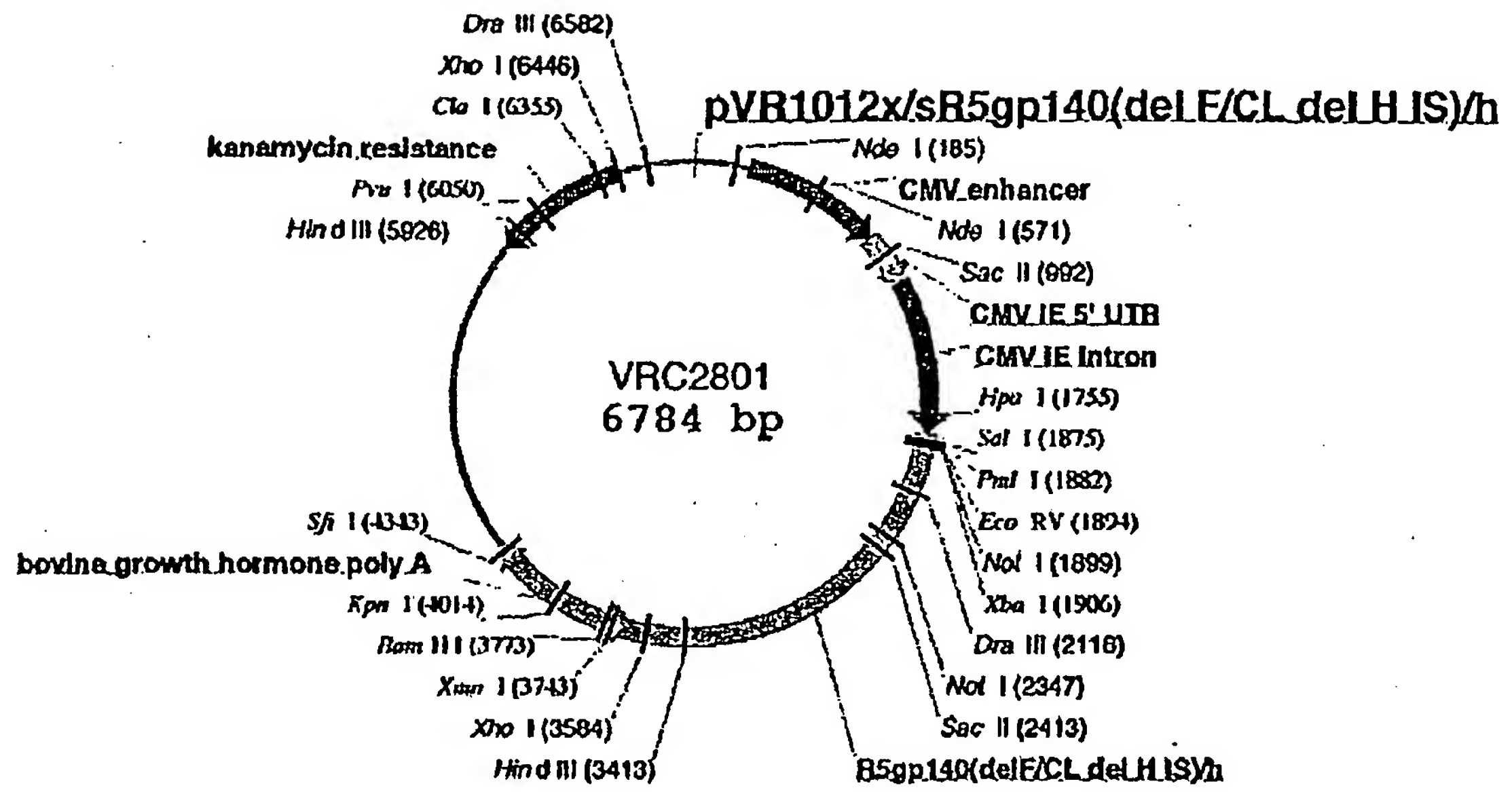


Figure 12



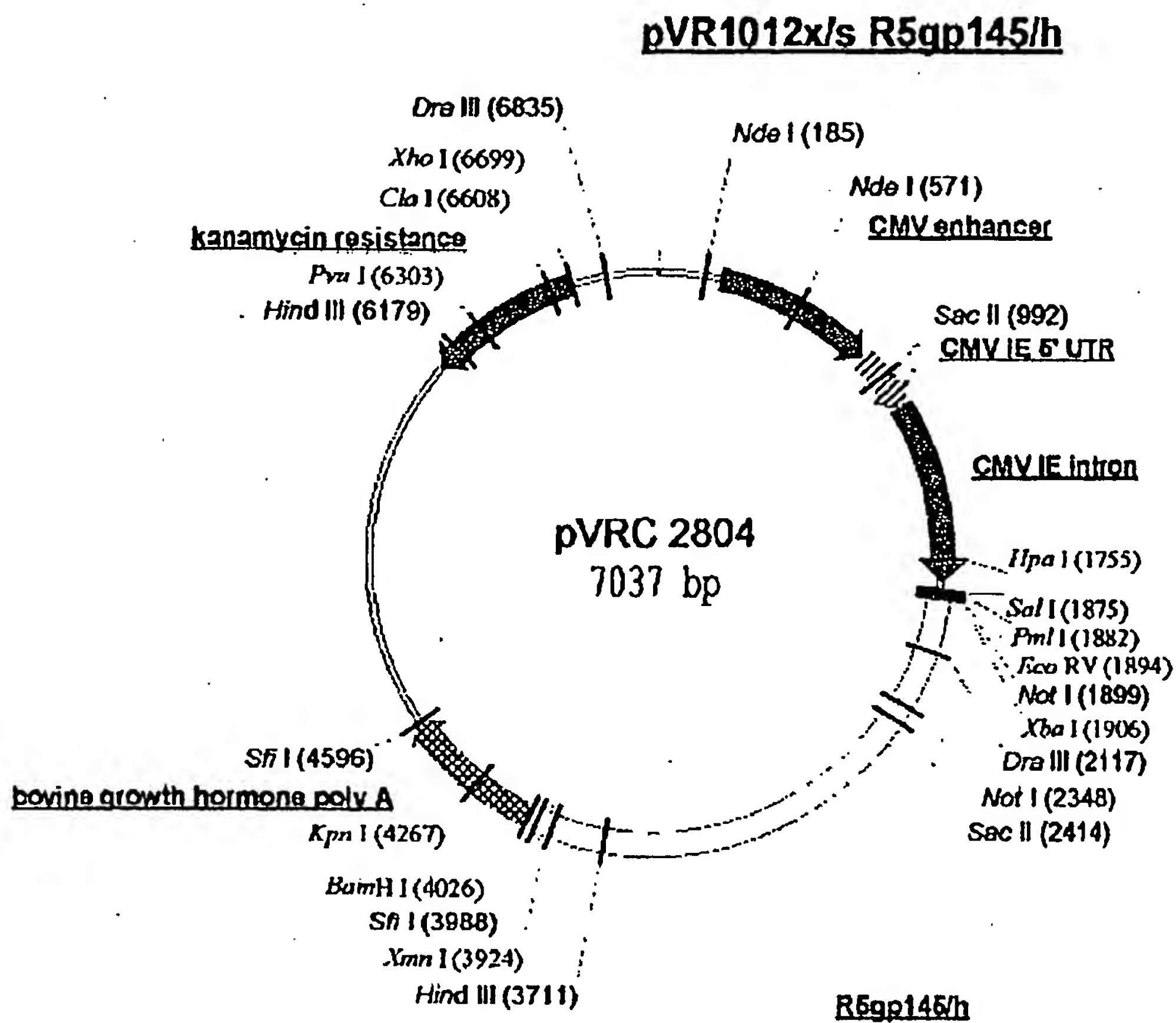


Figure 13

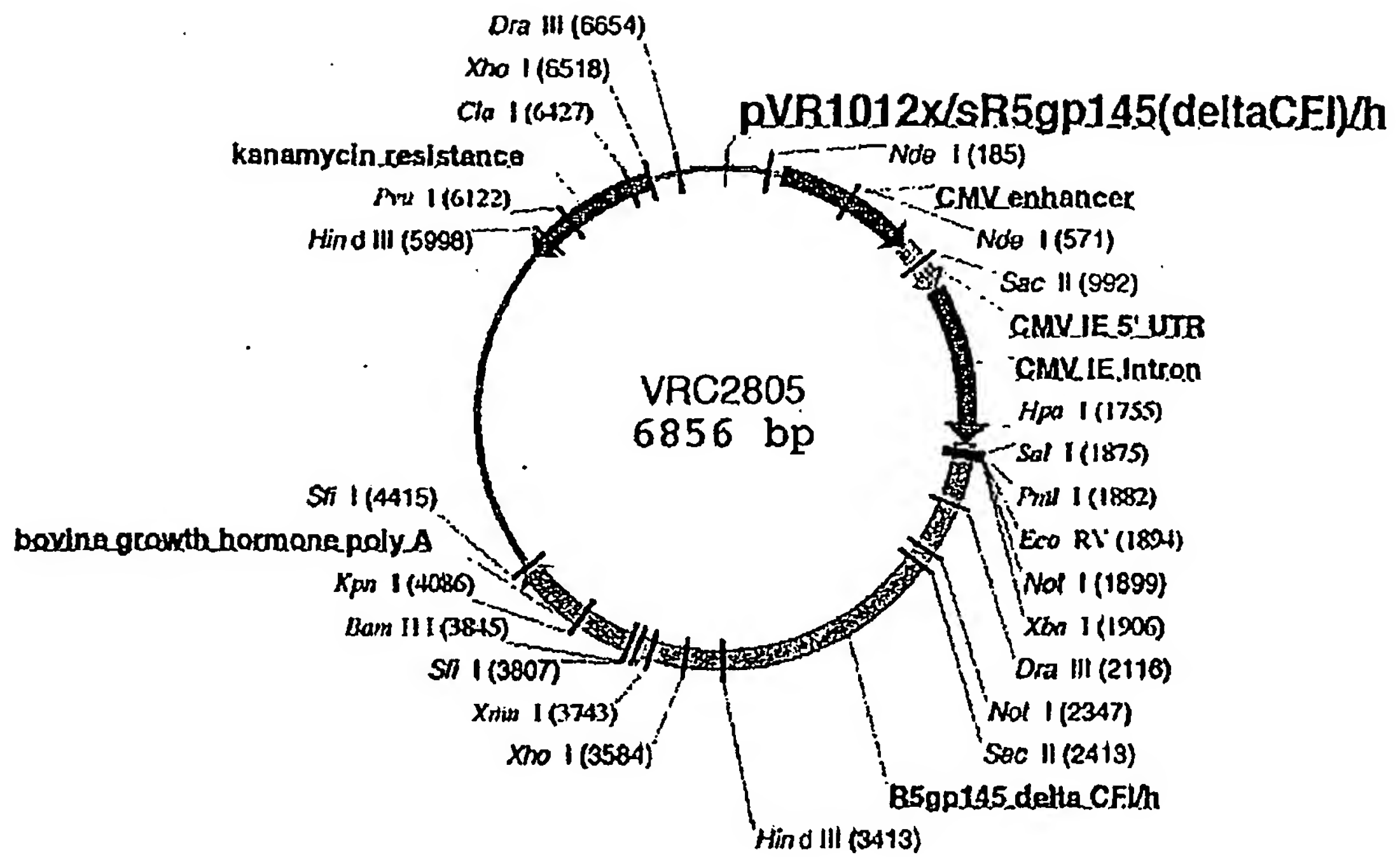


Figure 14

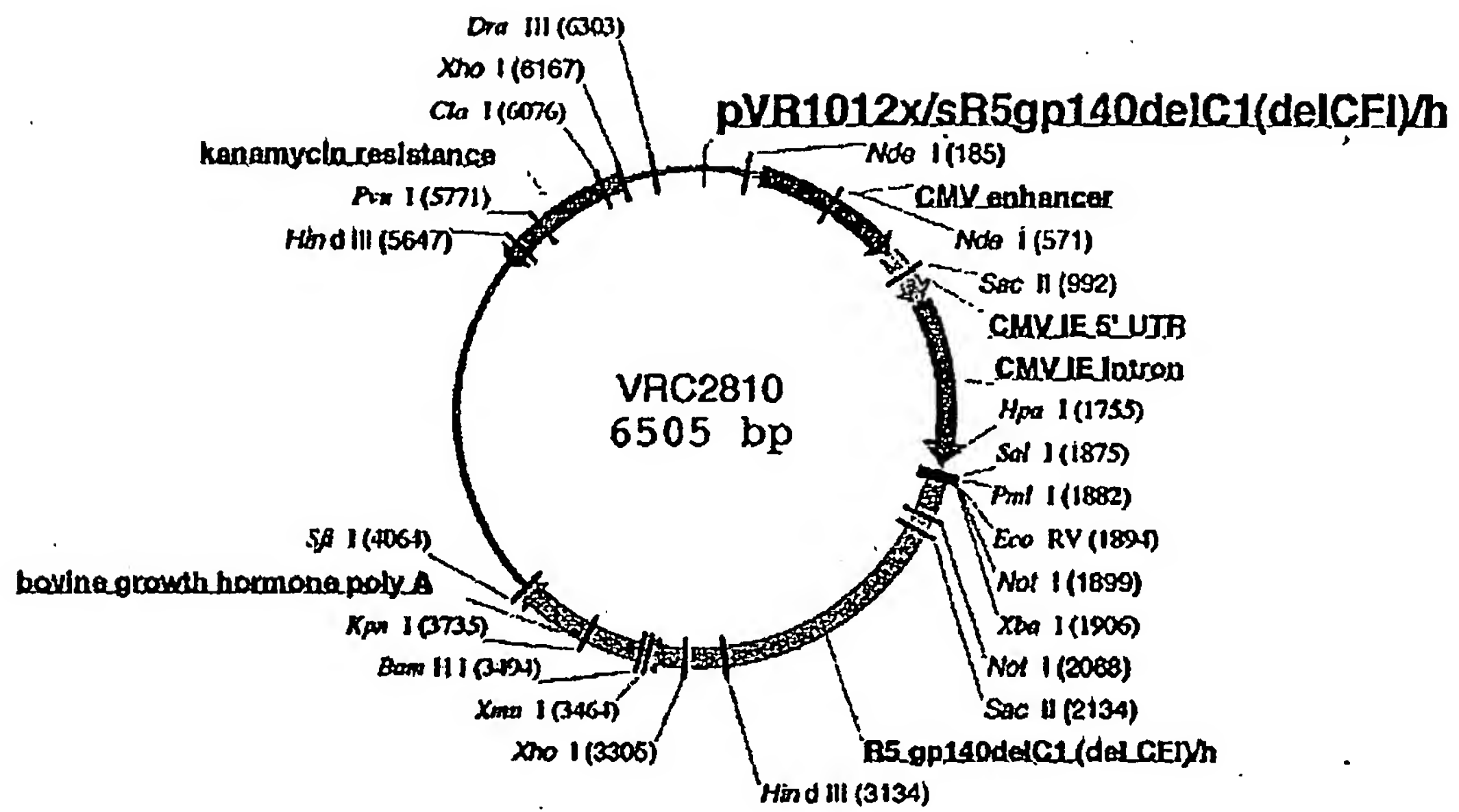


Figure 15

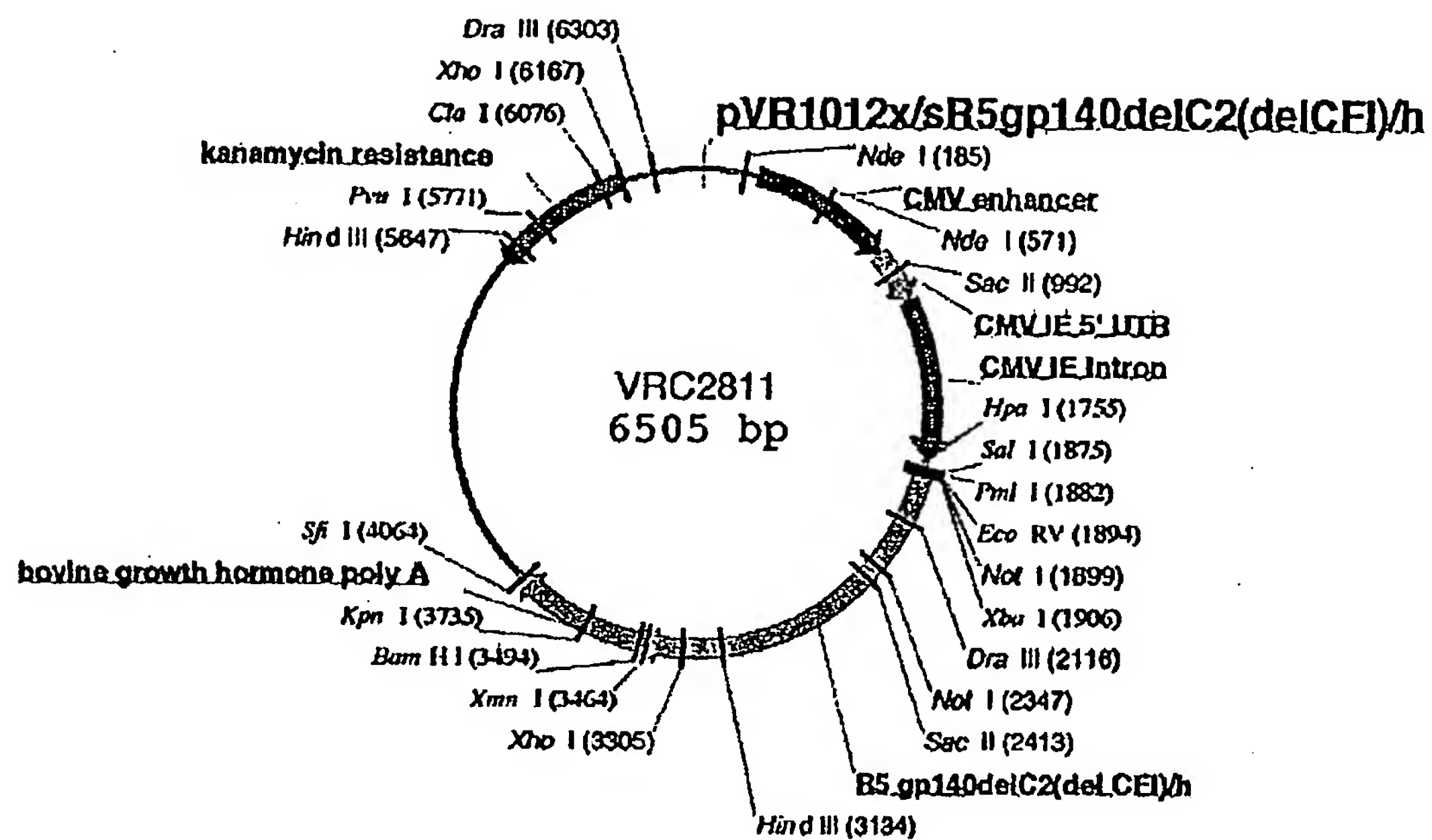


Figure 16

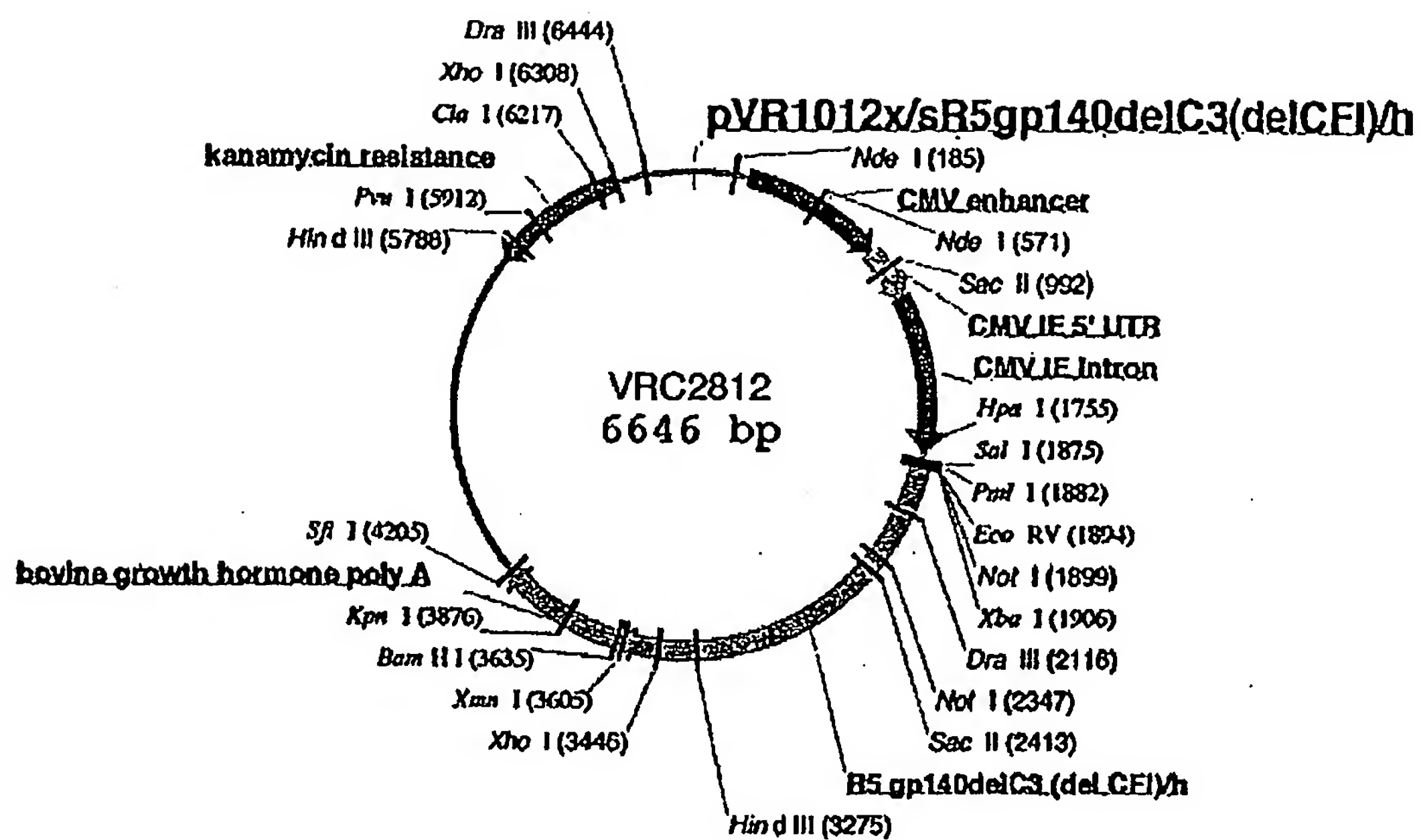


Figure 17



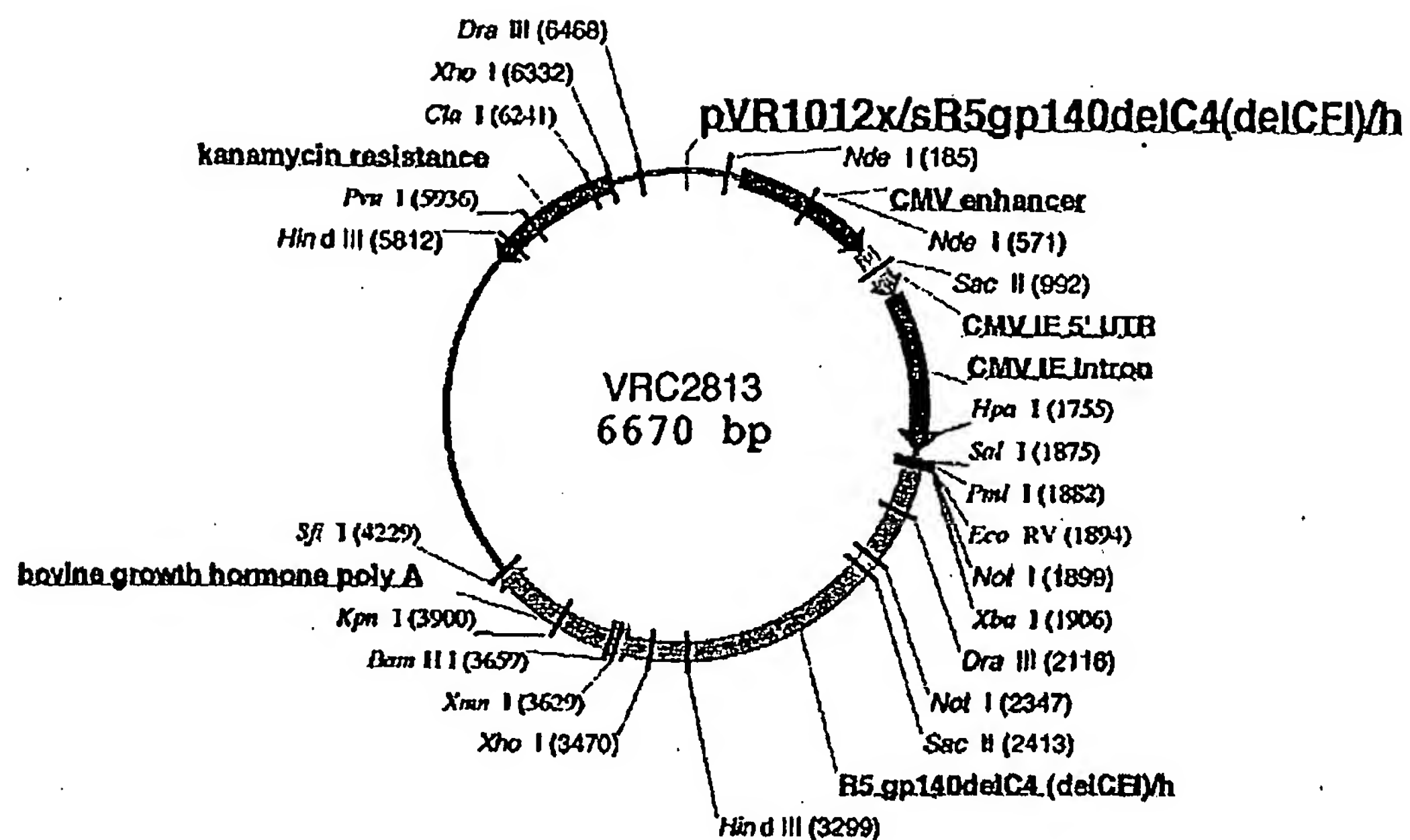


Figure 18

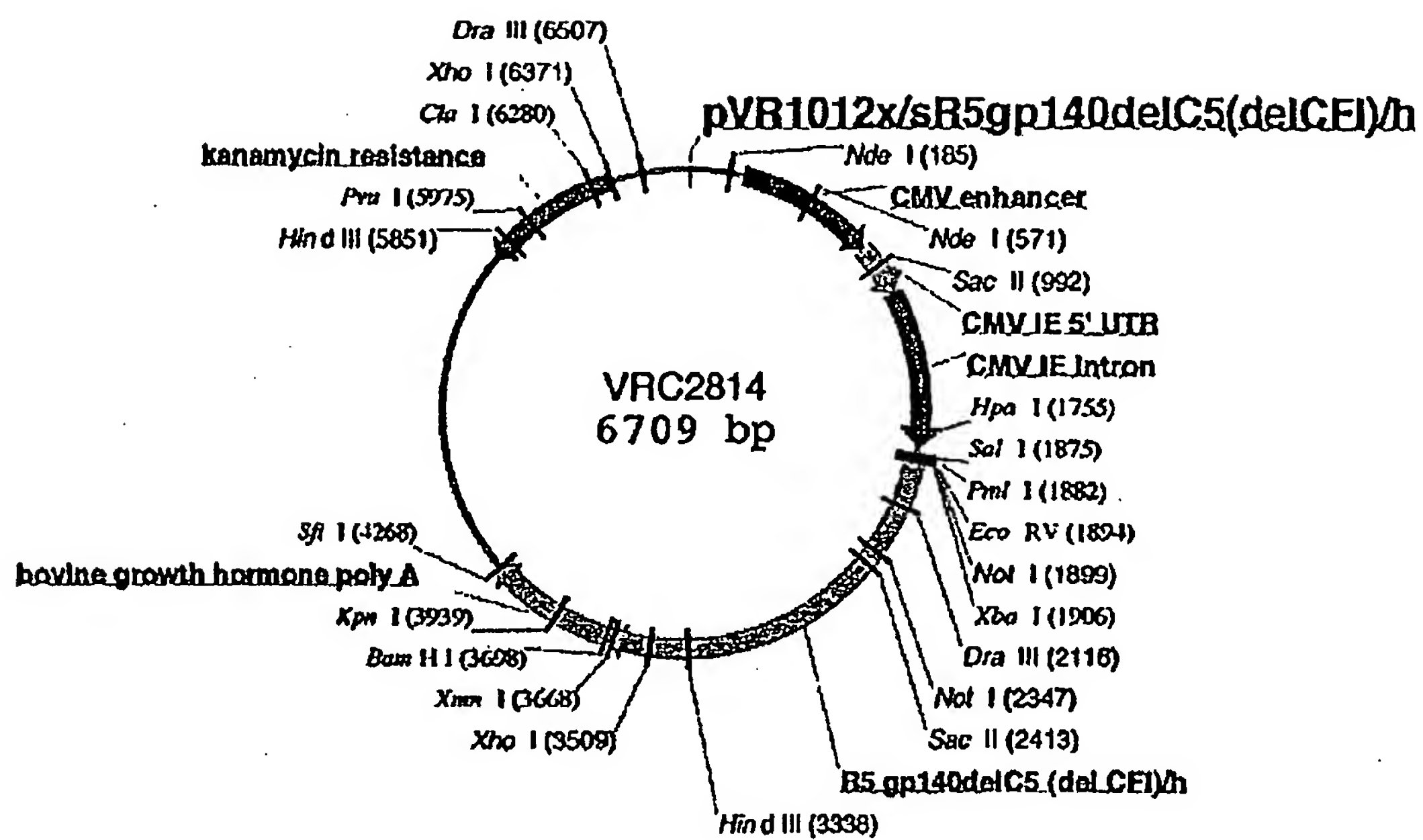


Figure 19

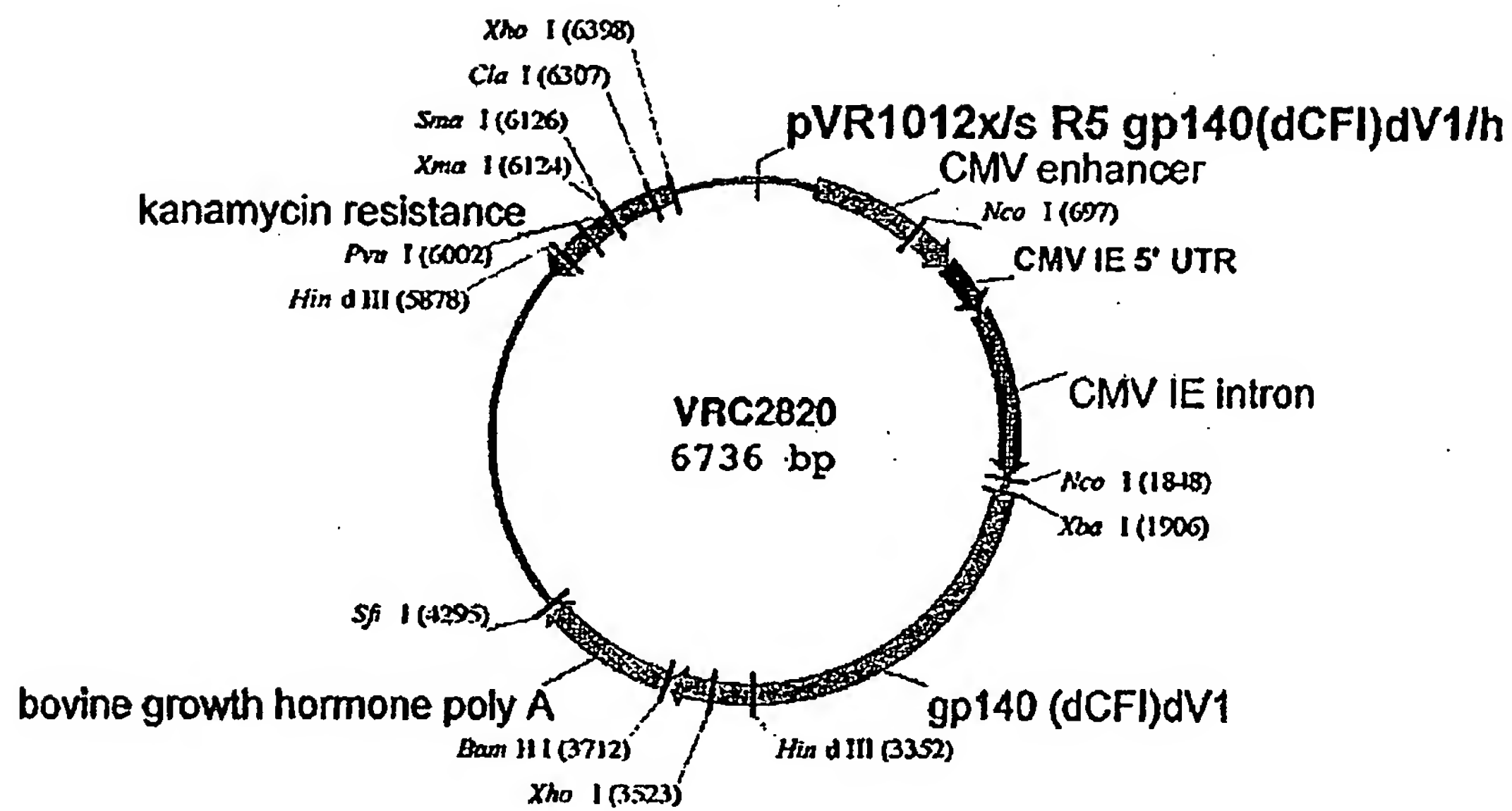


Figure 20

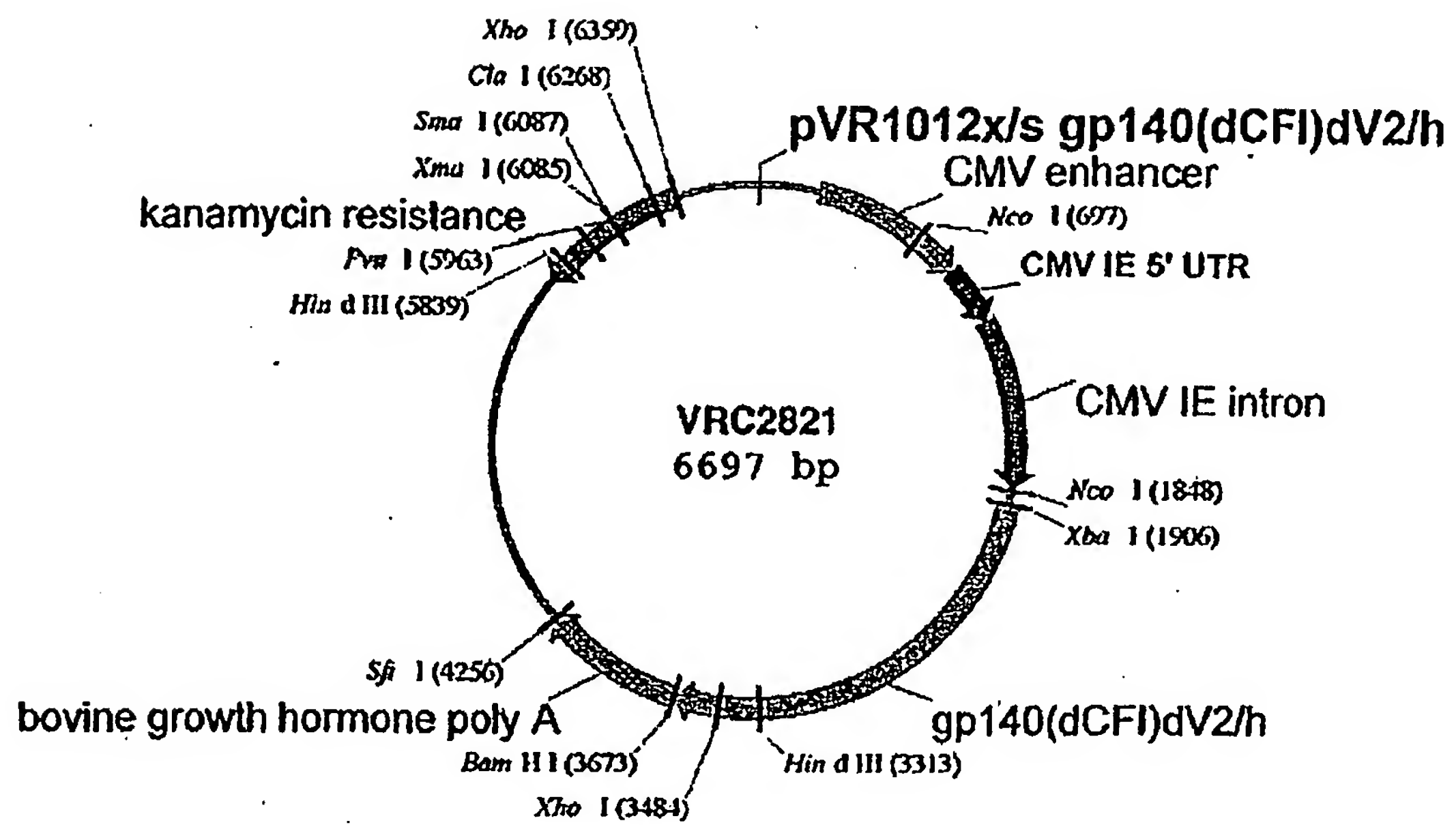


Figure 21

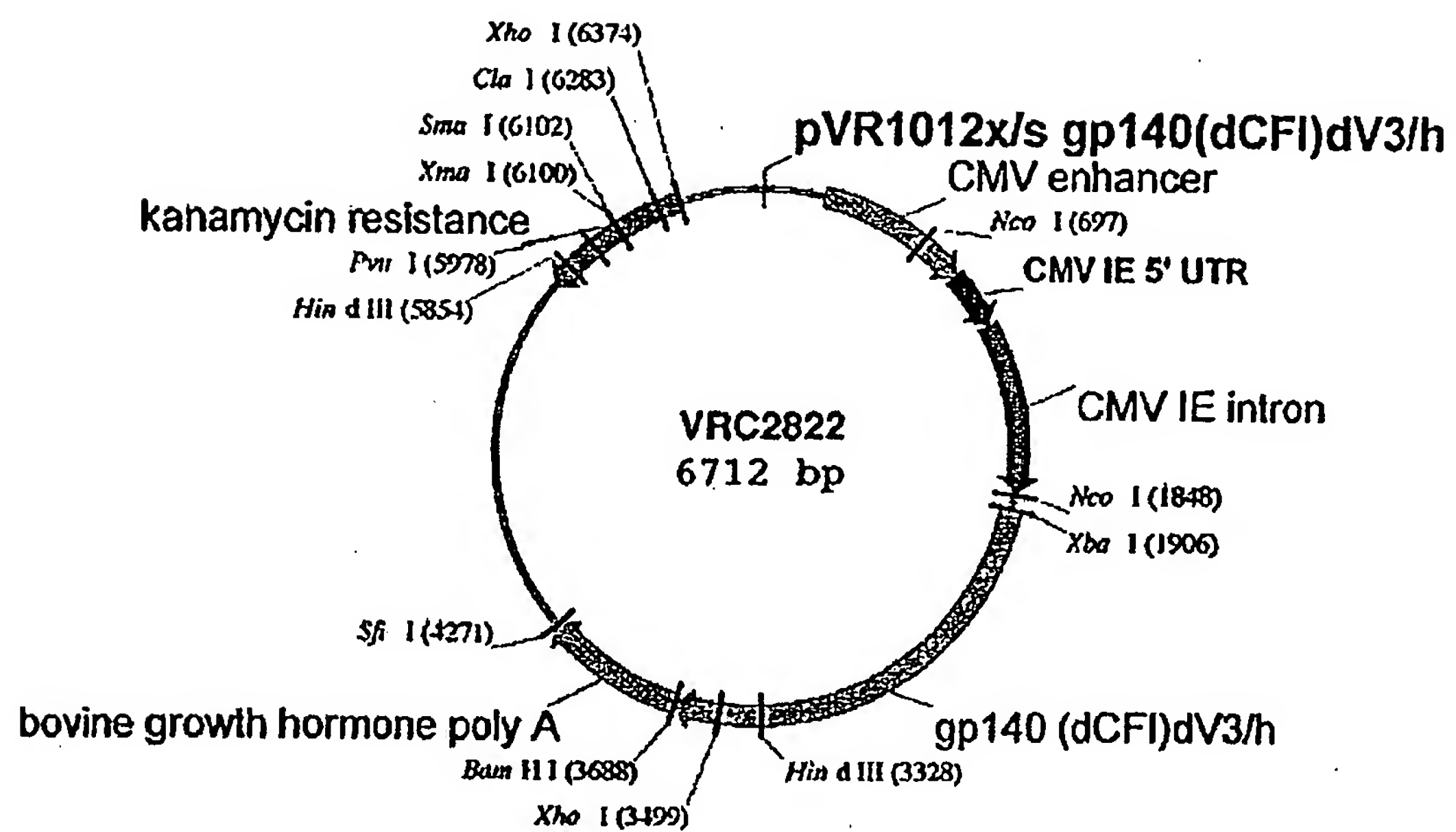


Figure 22

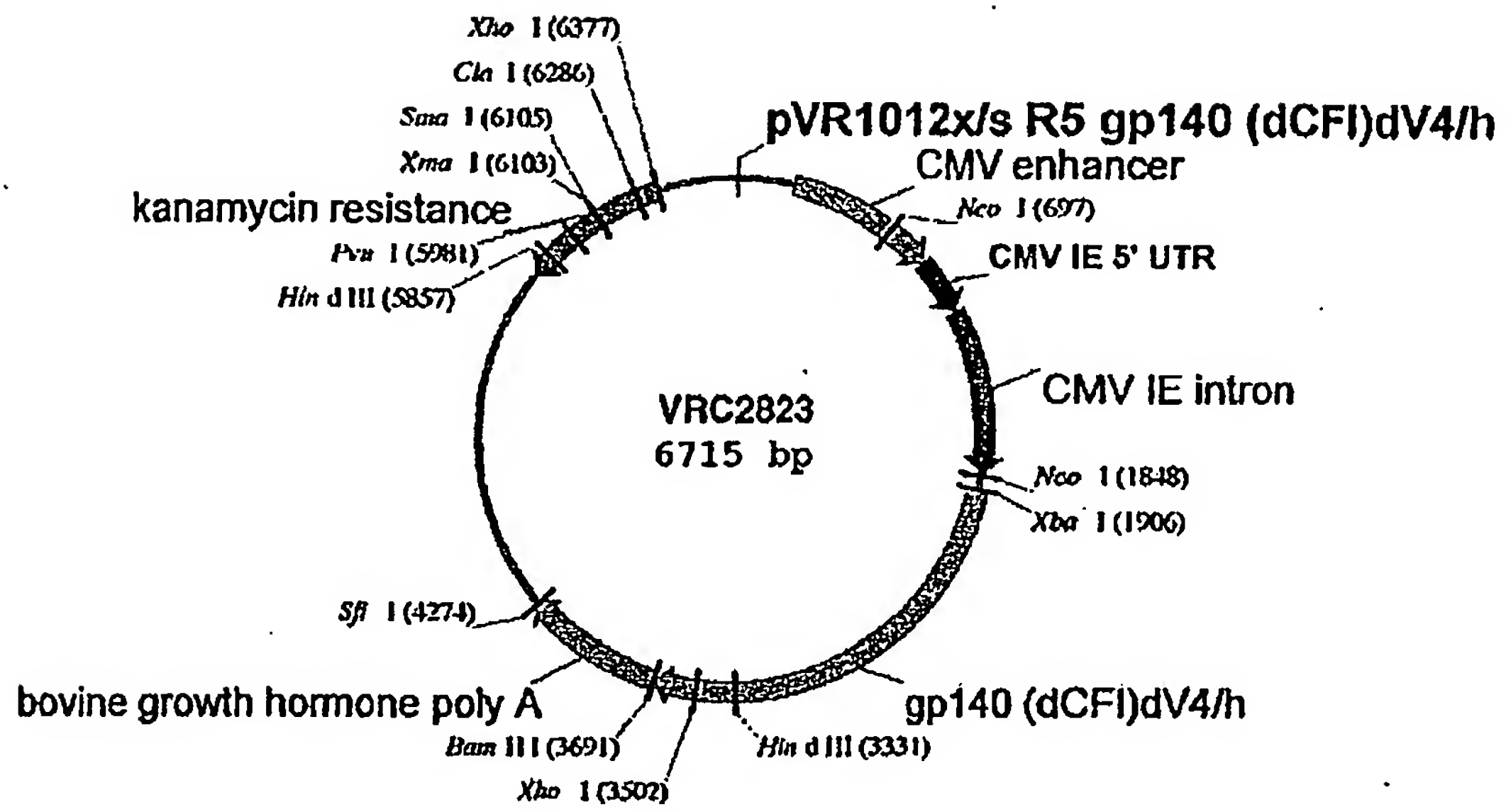


Figure 23



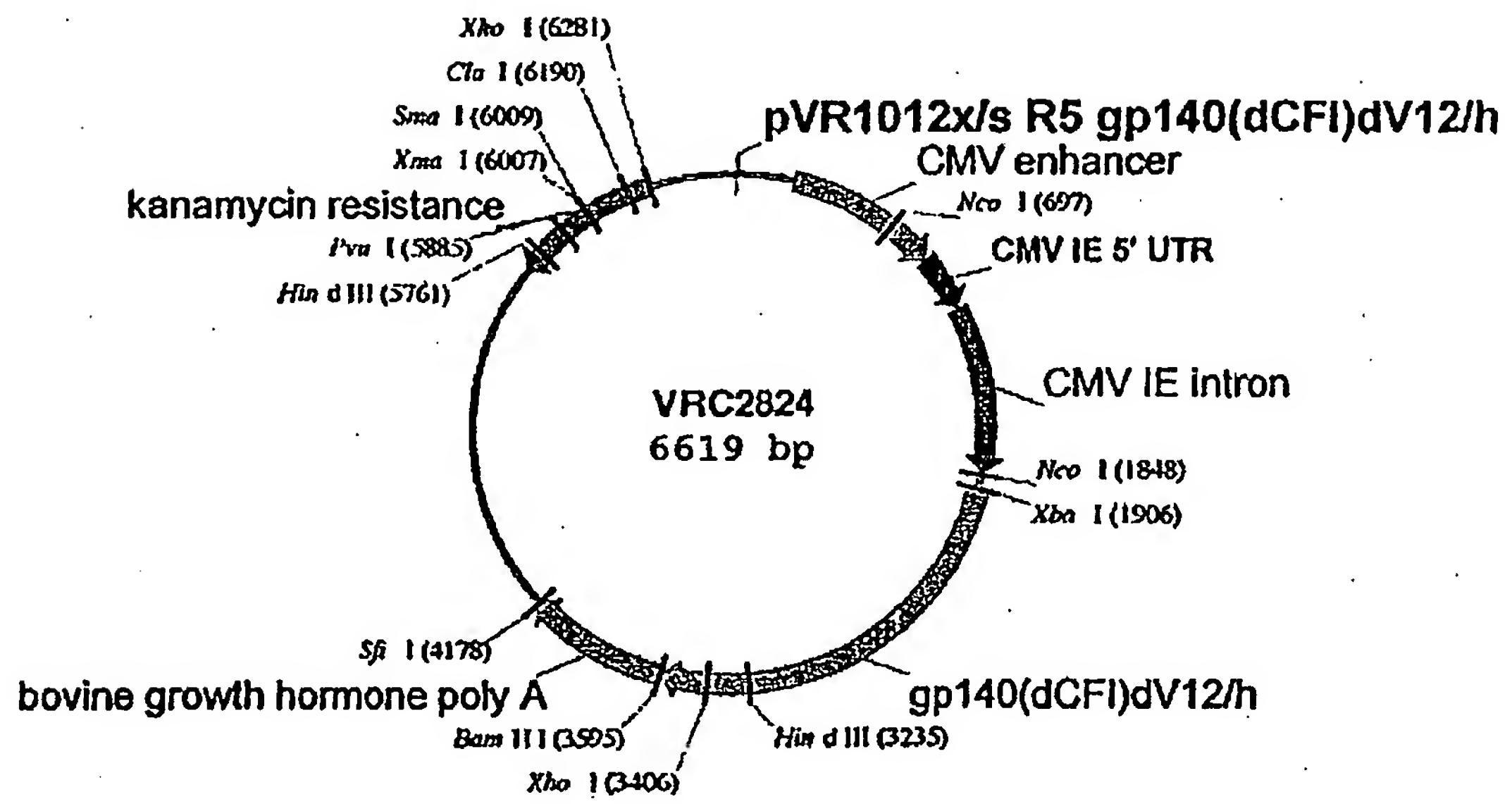


Figure 24

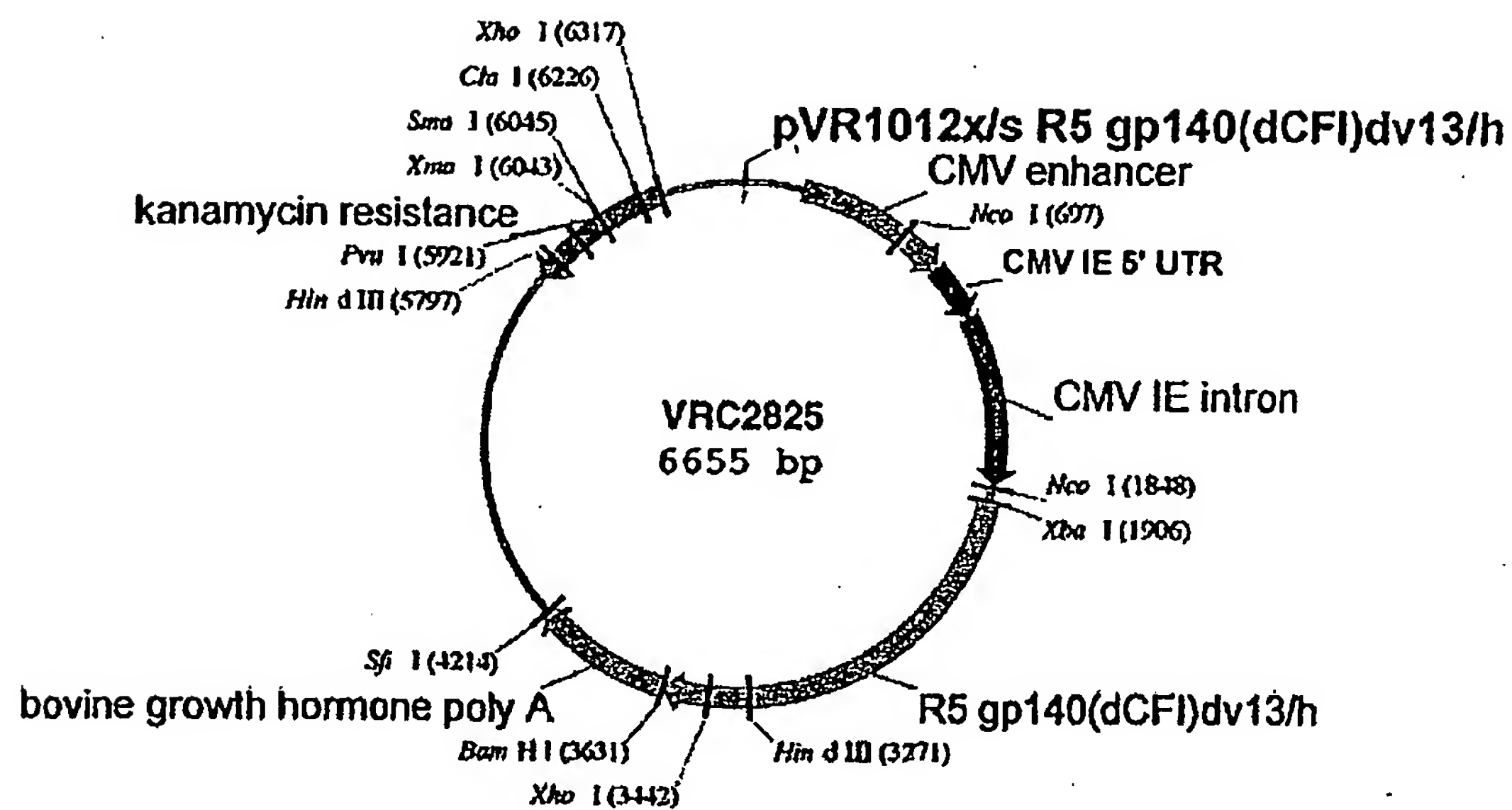


Figure 25

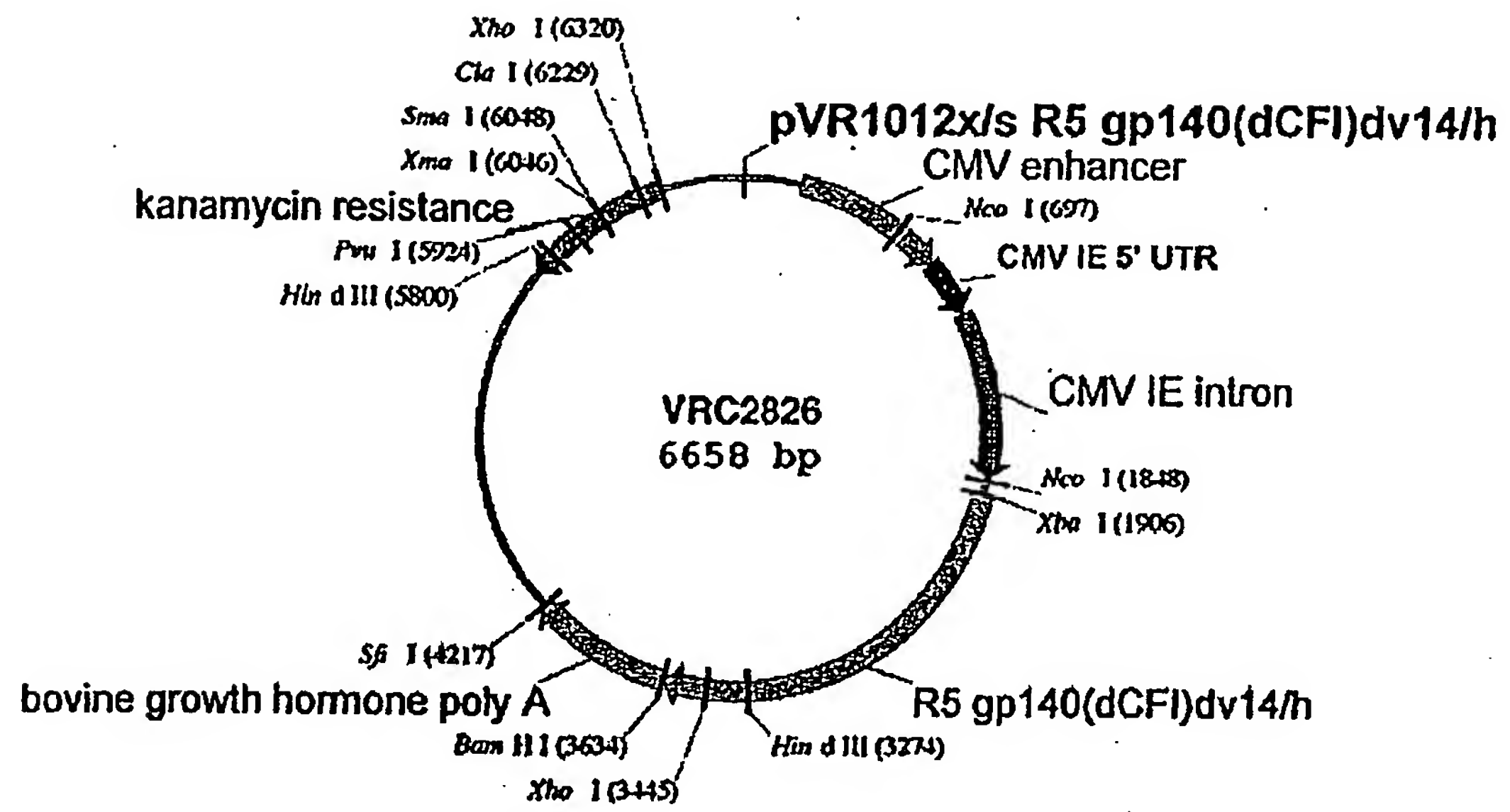


Figure 26

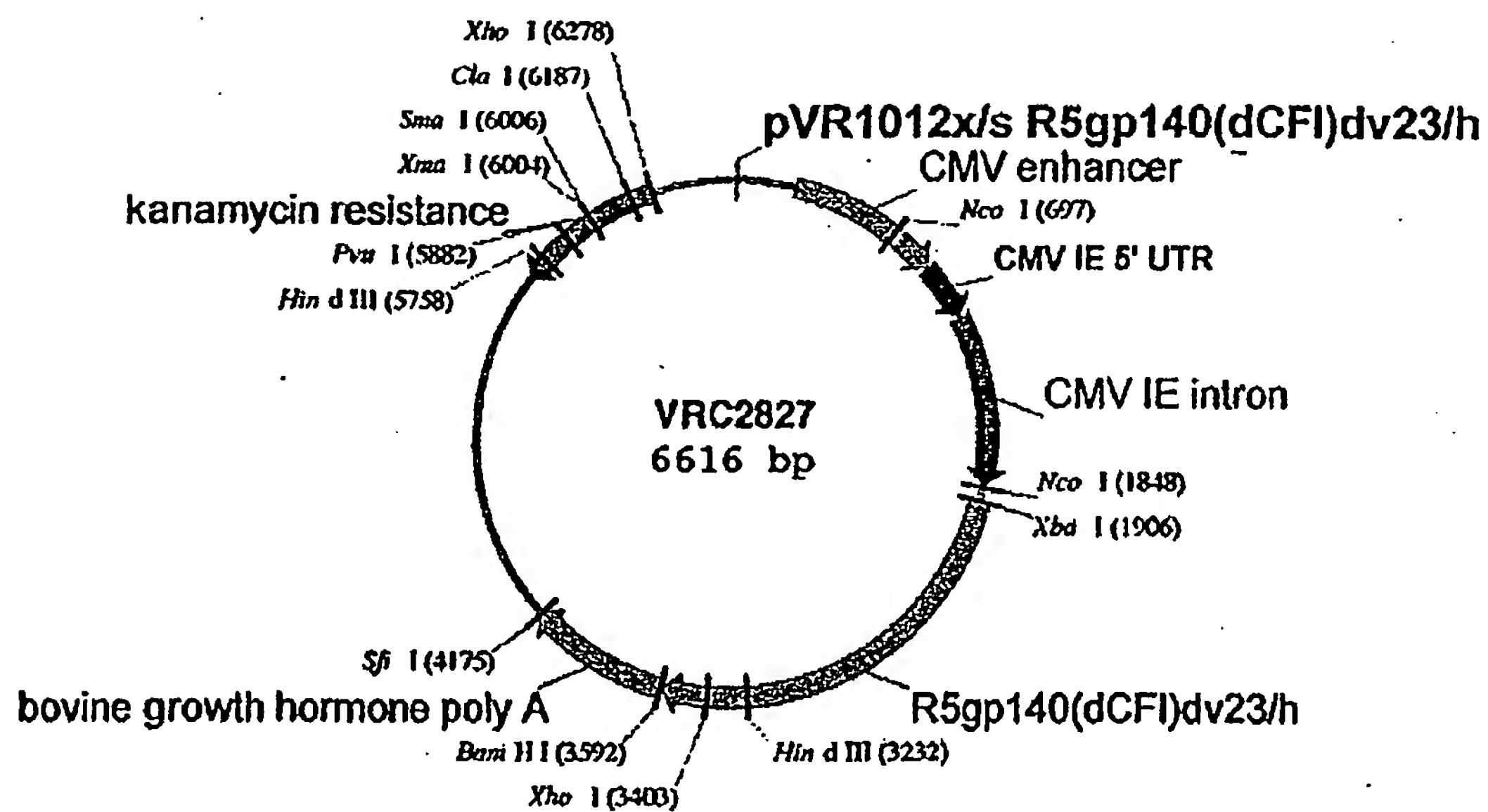


Figure 27

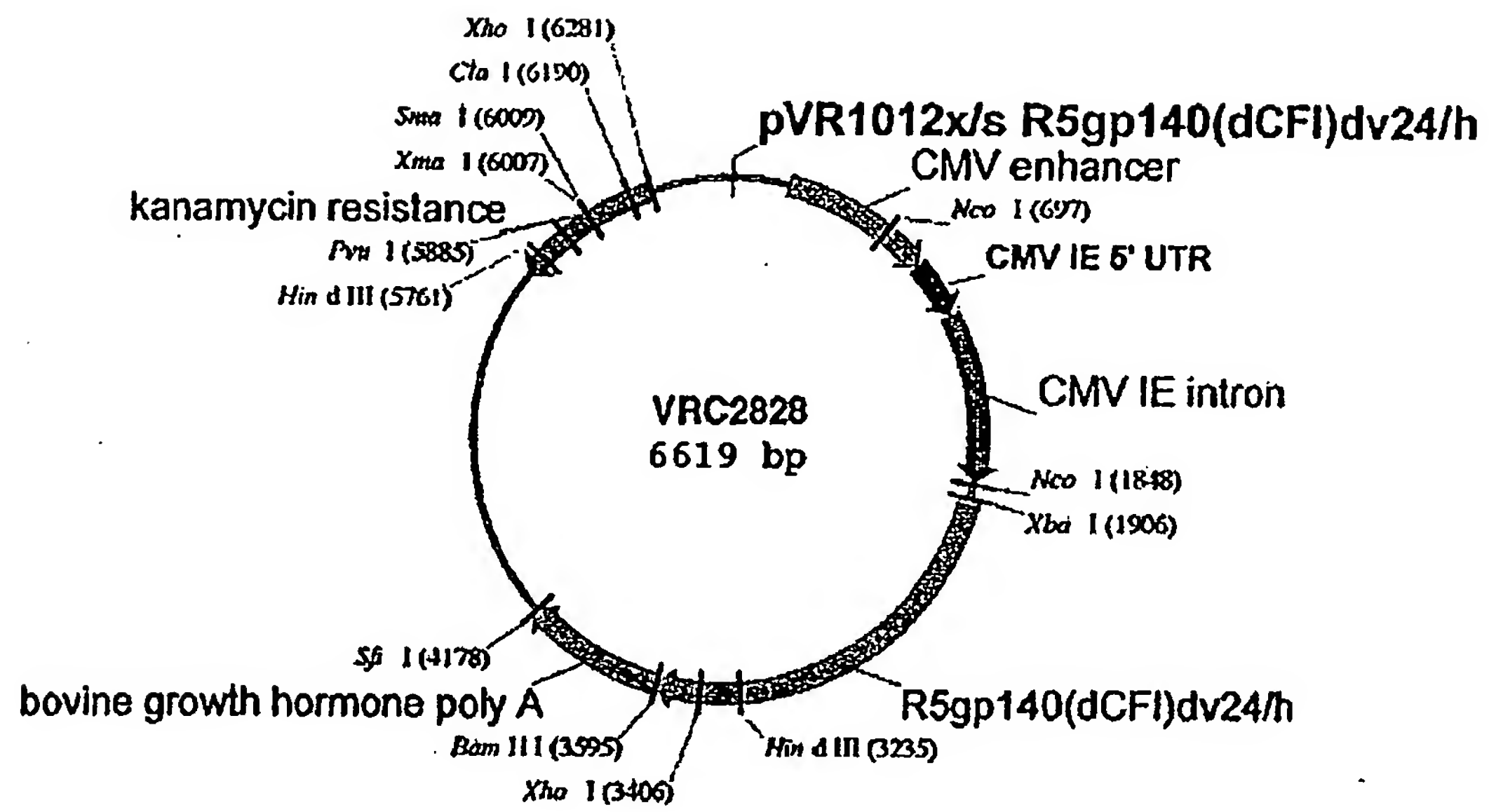


Figure 28

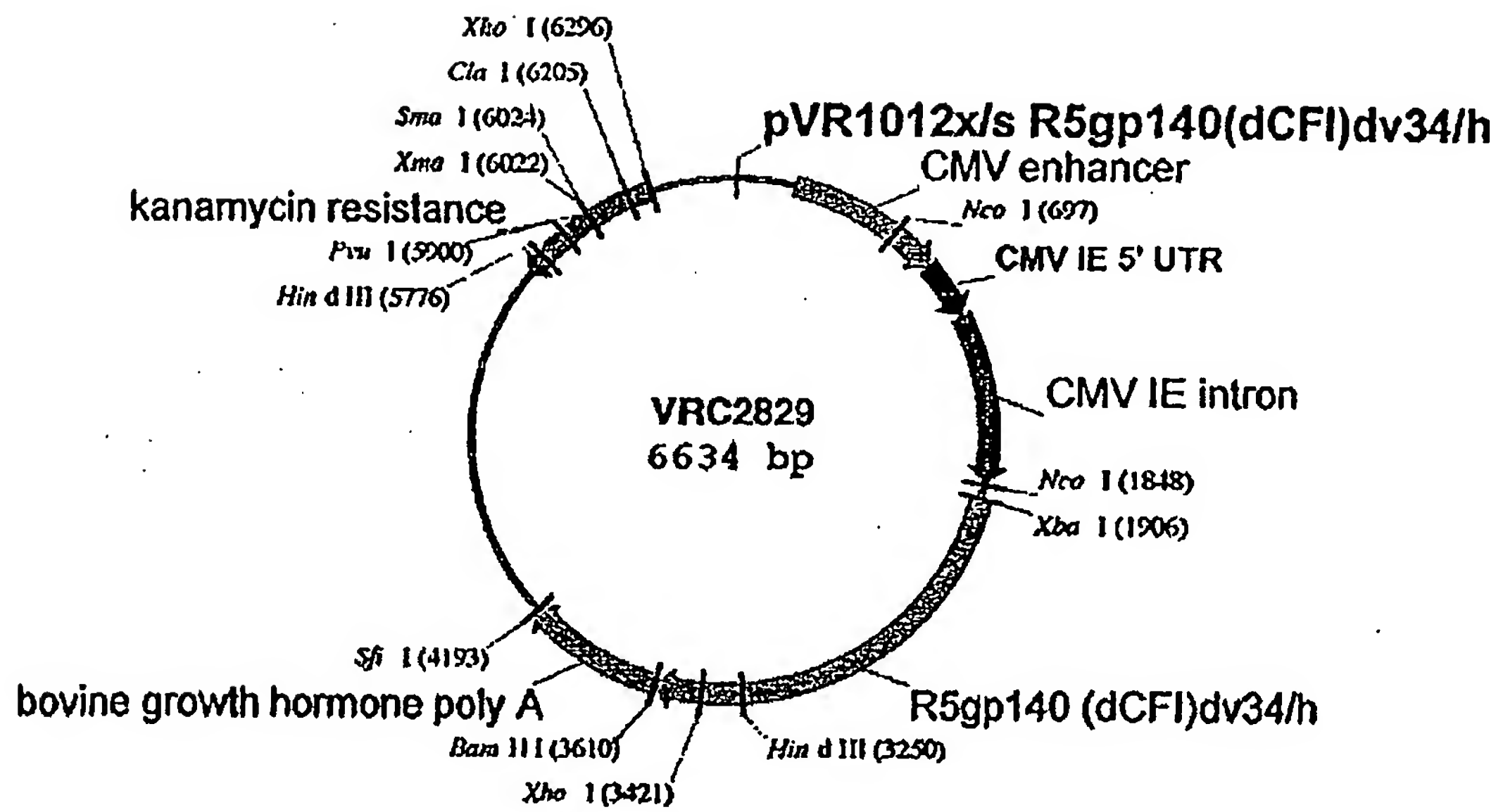


Figure 29



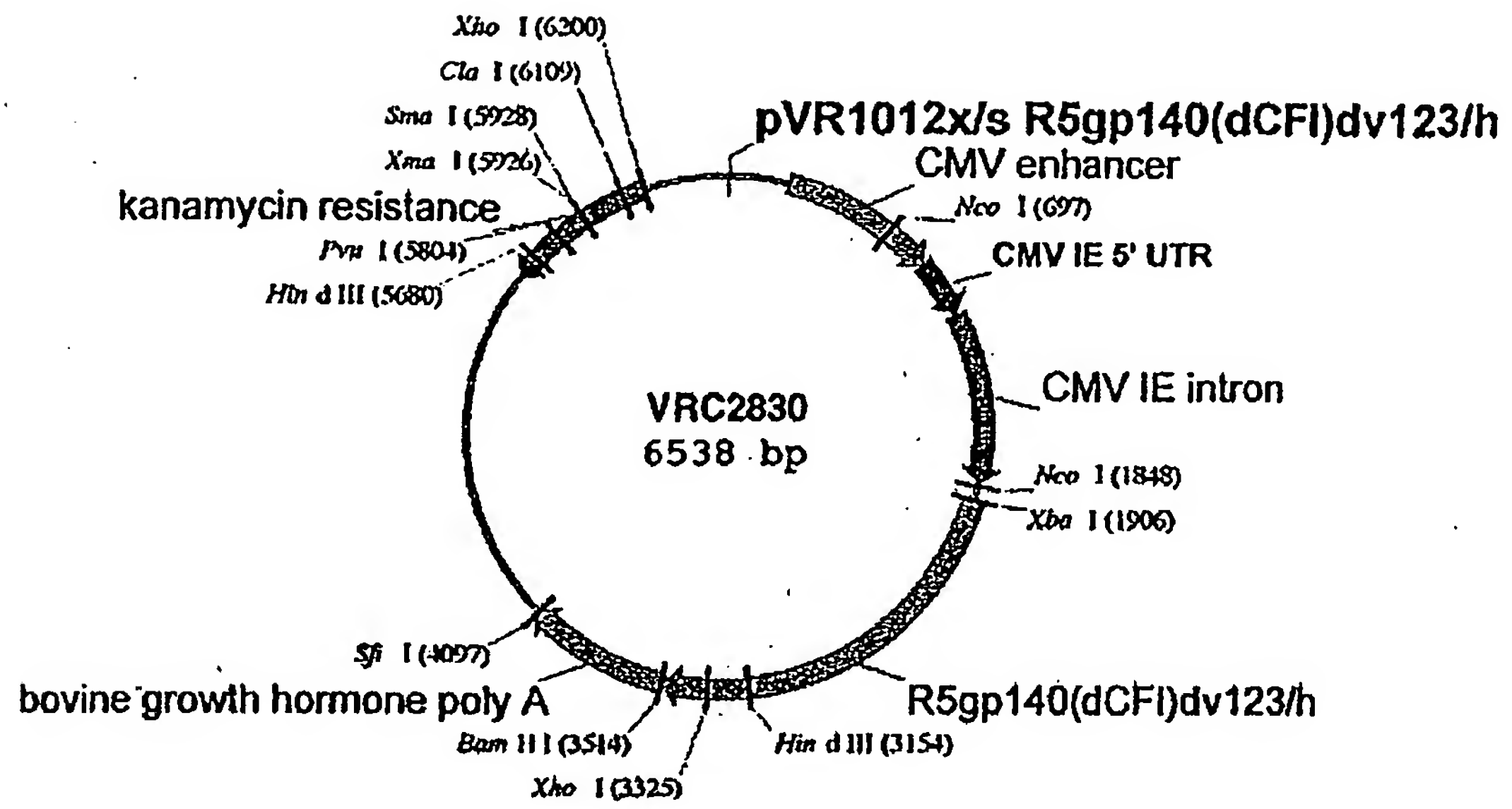


Figure 30

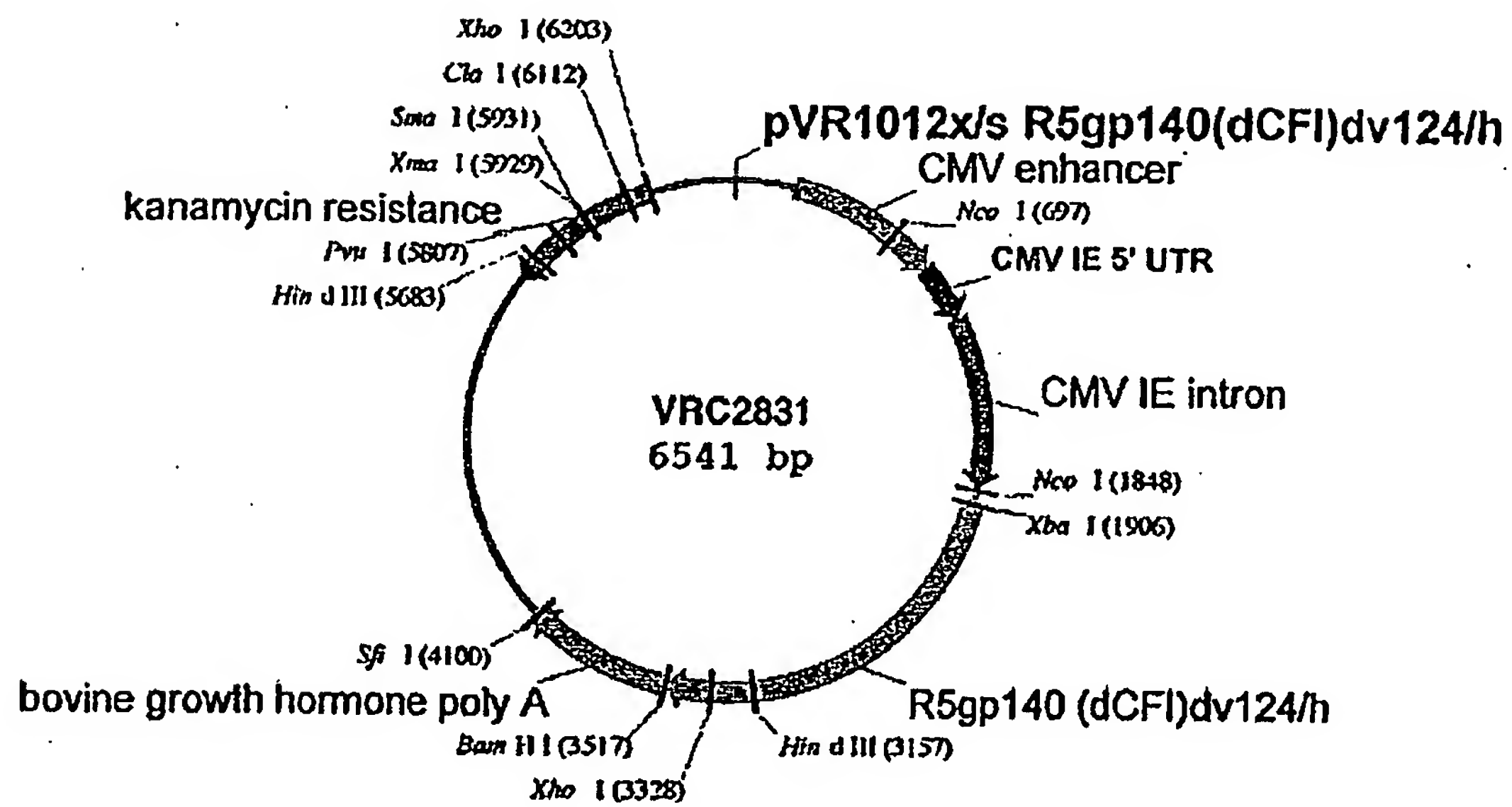


Figure 31

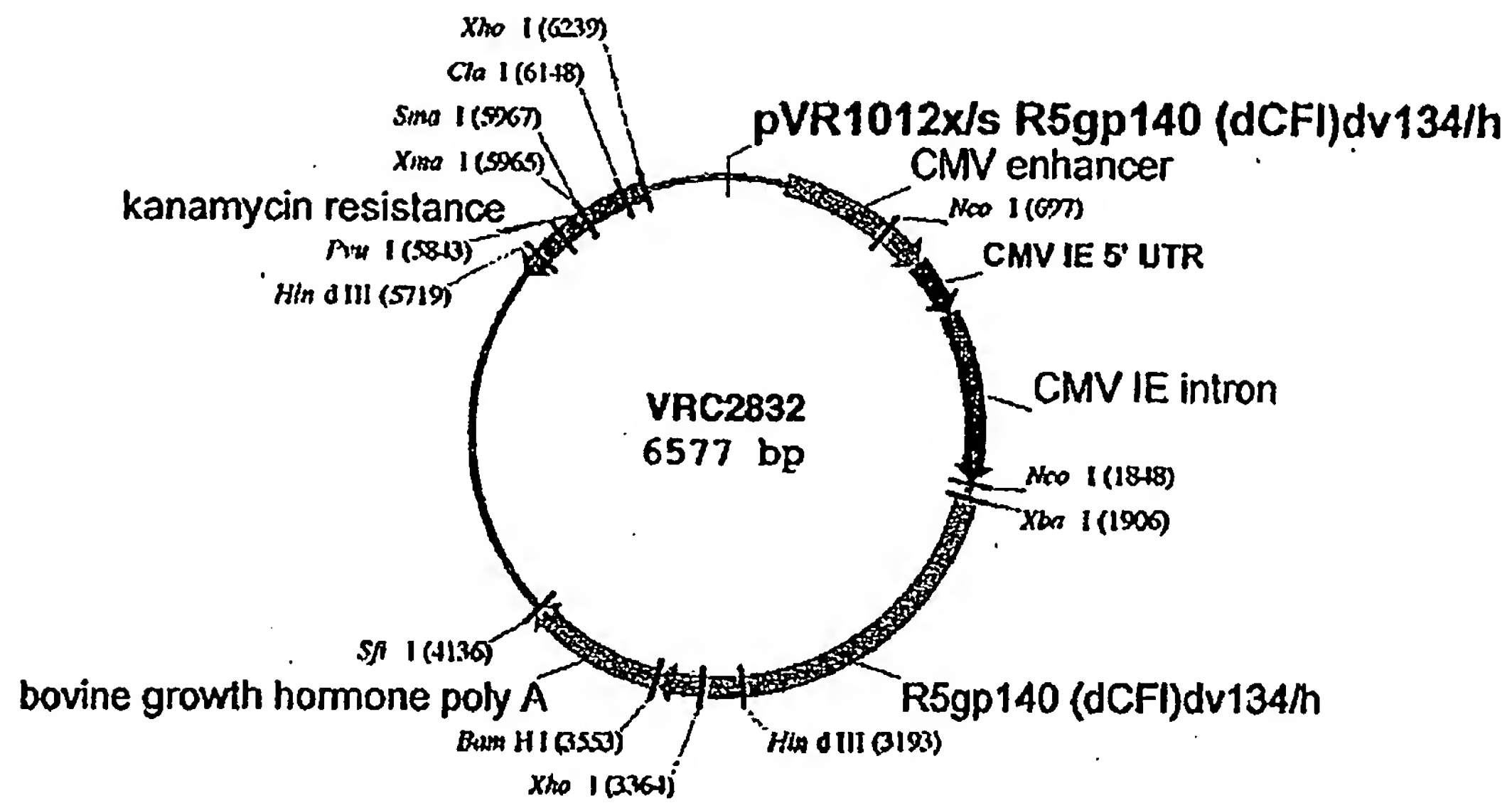


Figure 32

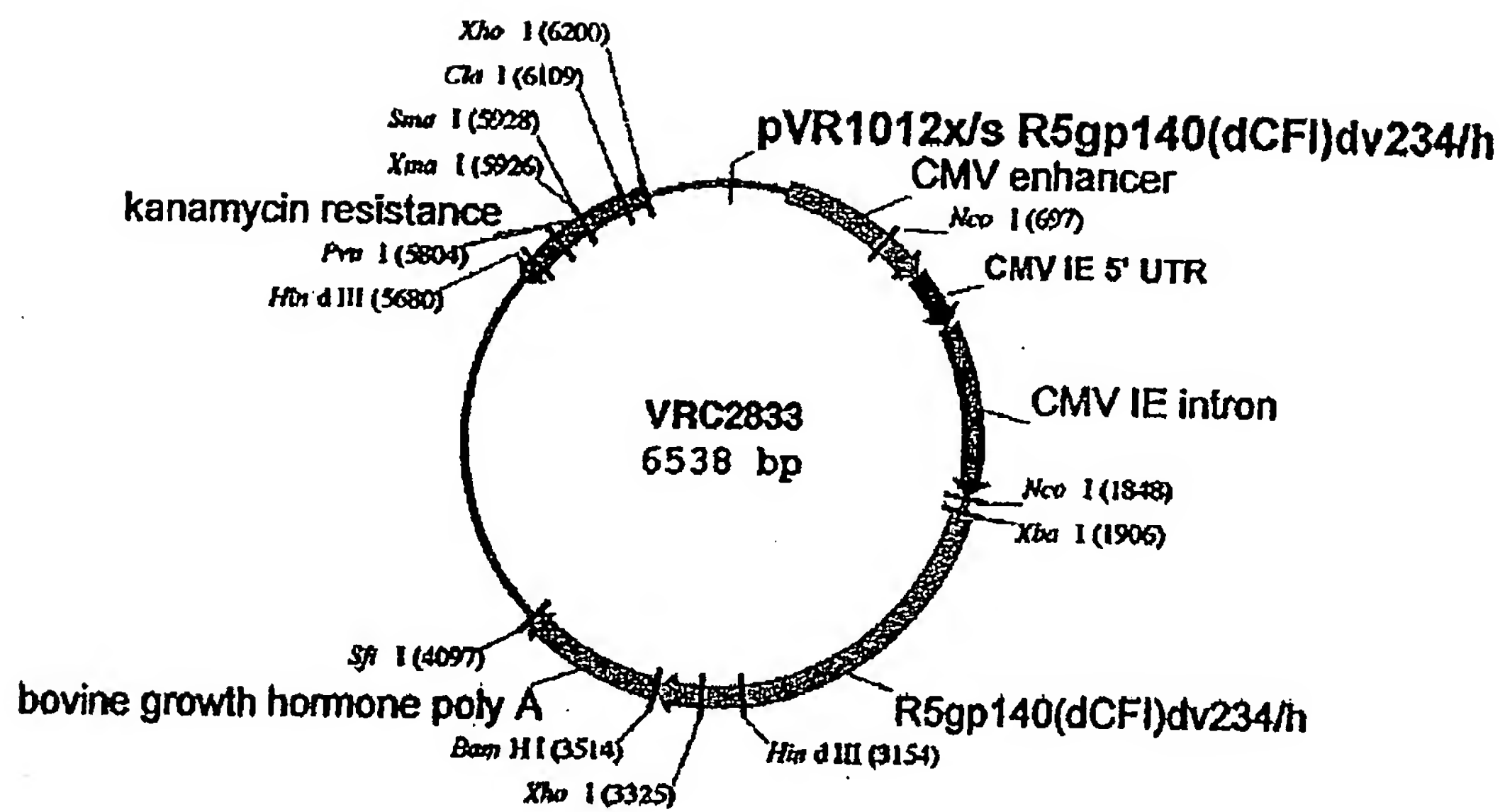


Figure 33

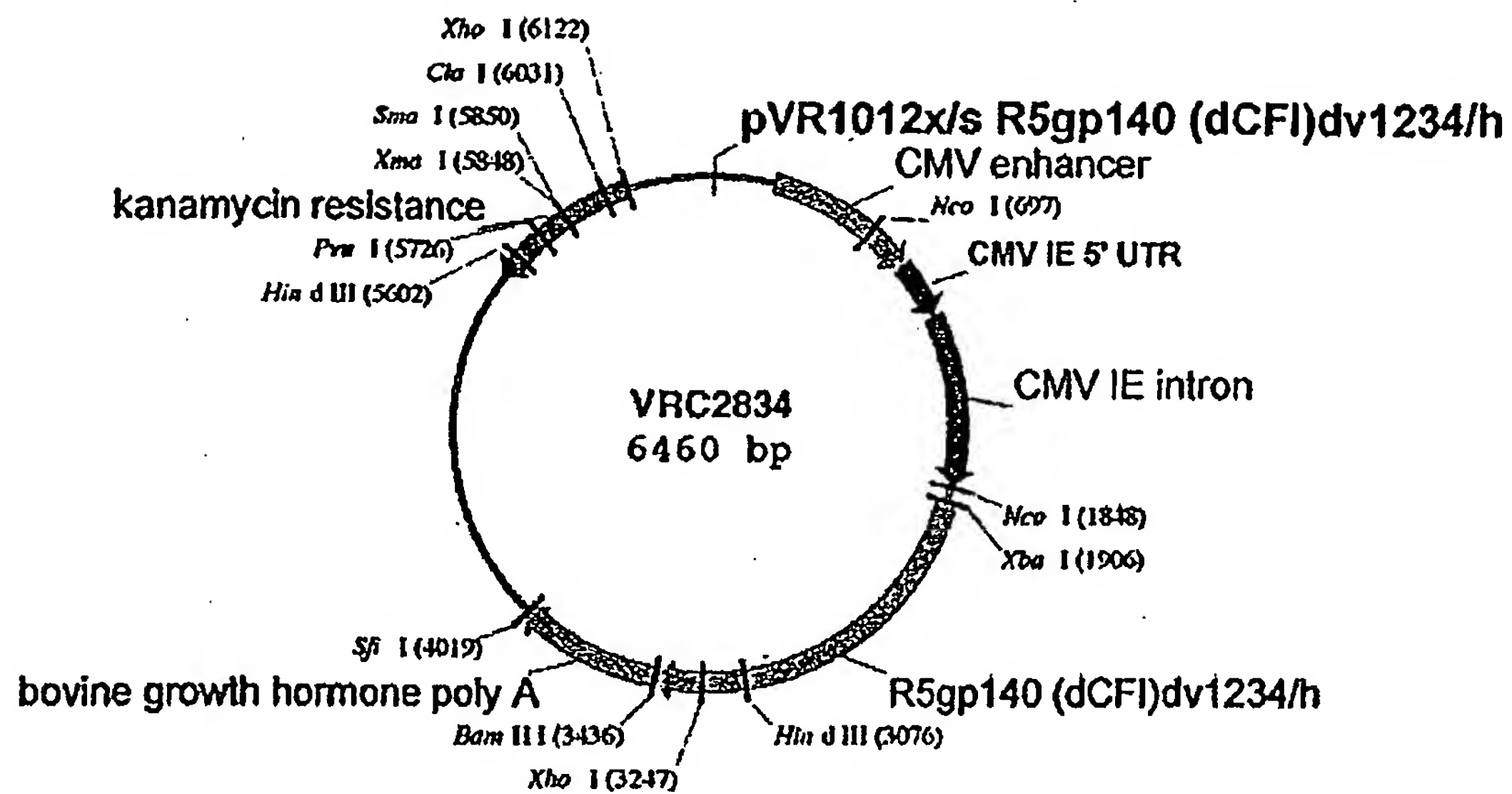


Figure 34

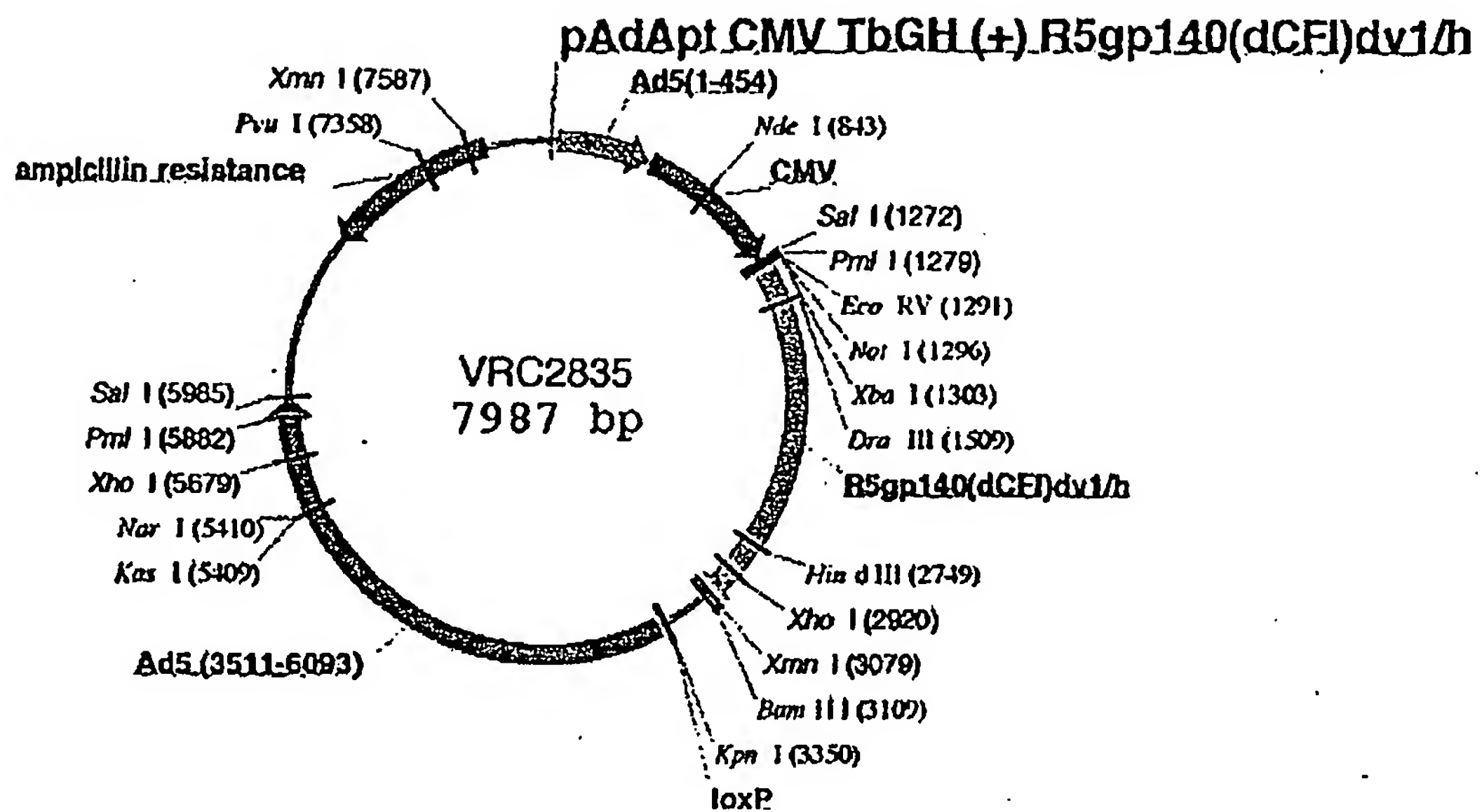


Figure 35



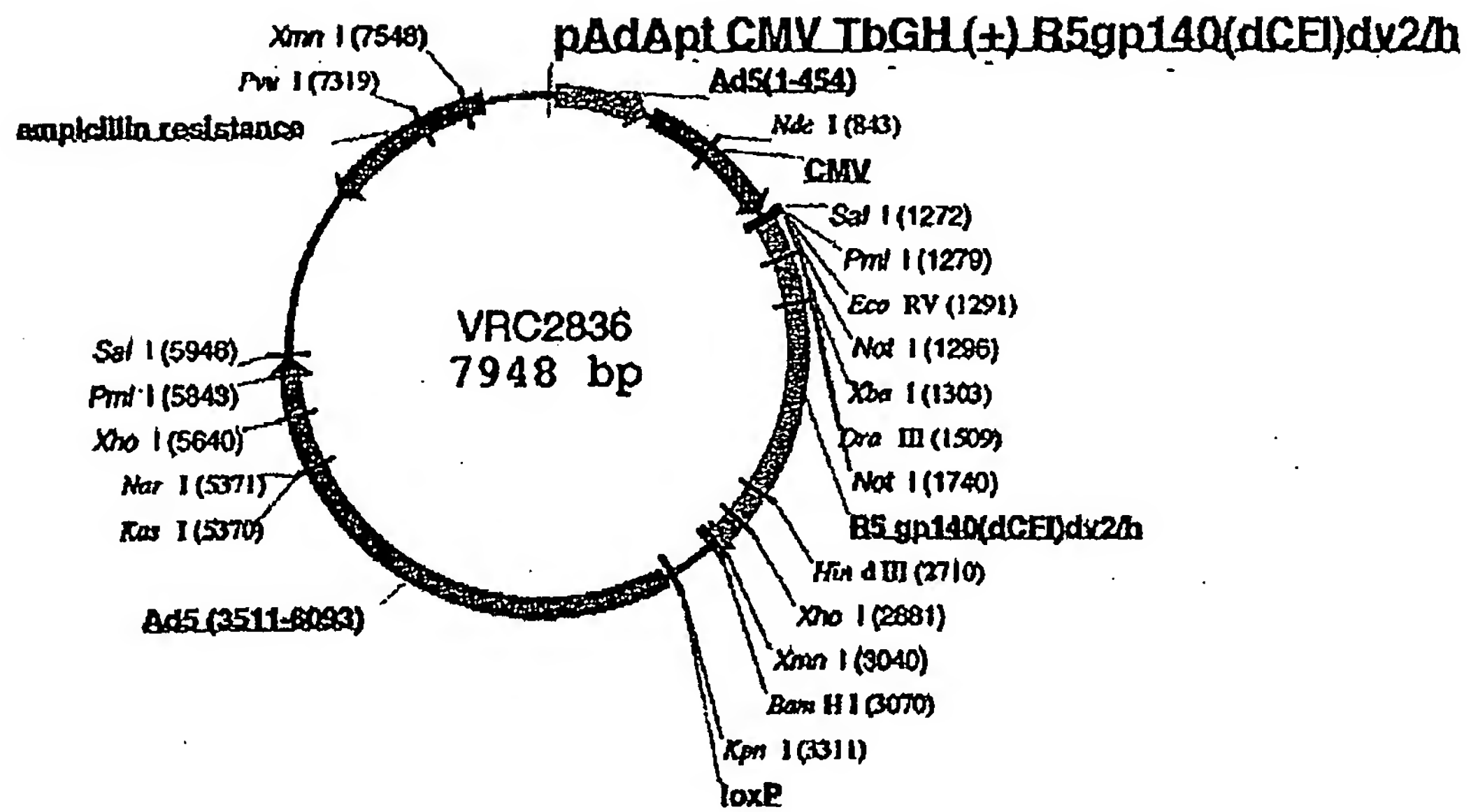


Figure 36

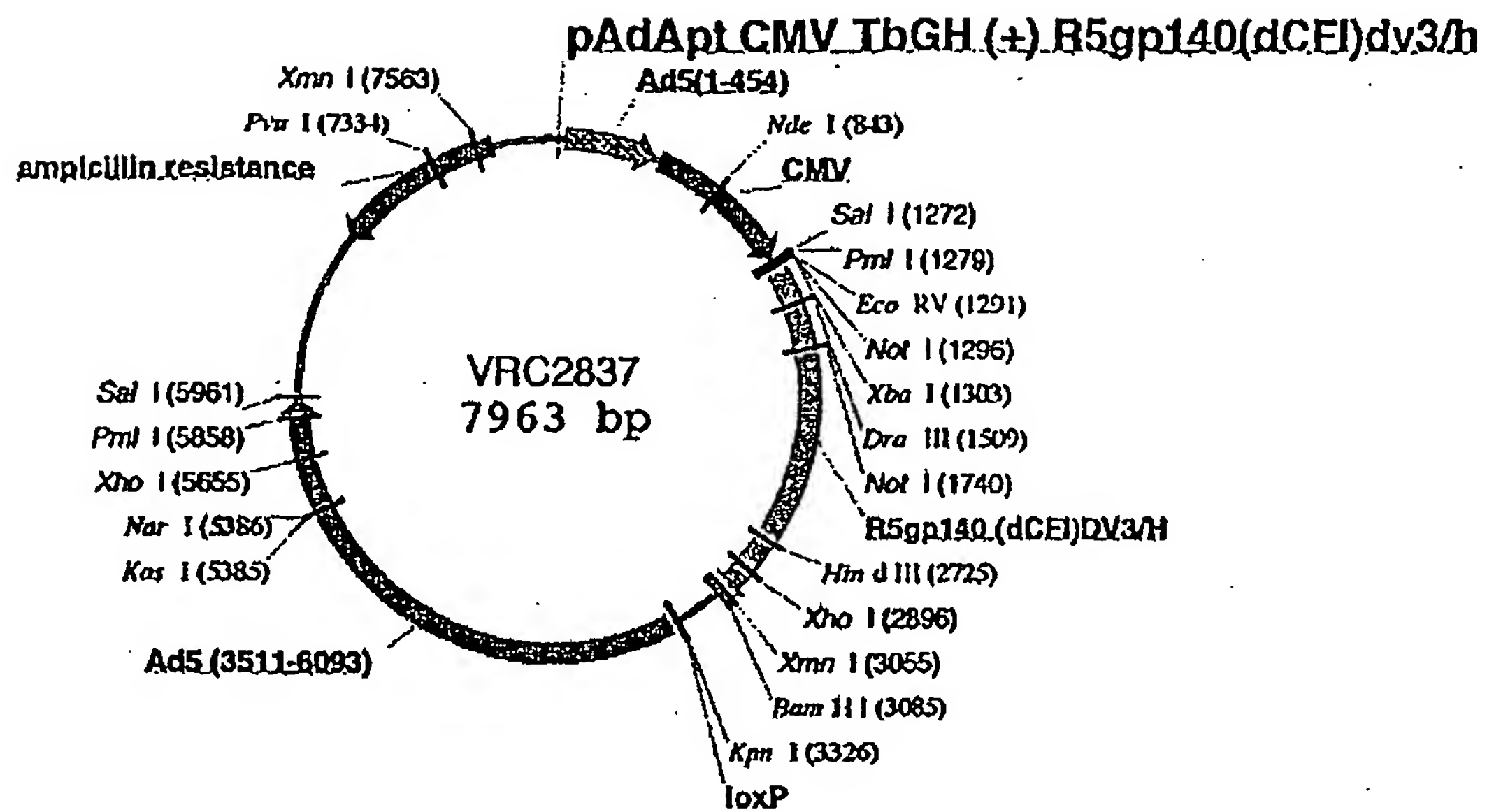


Figure 37

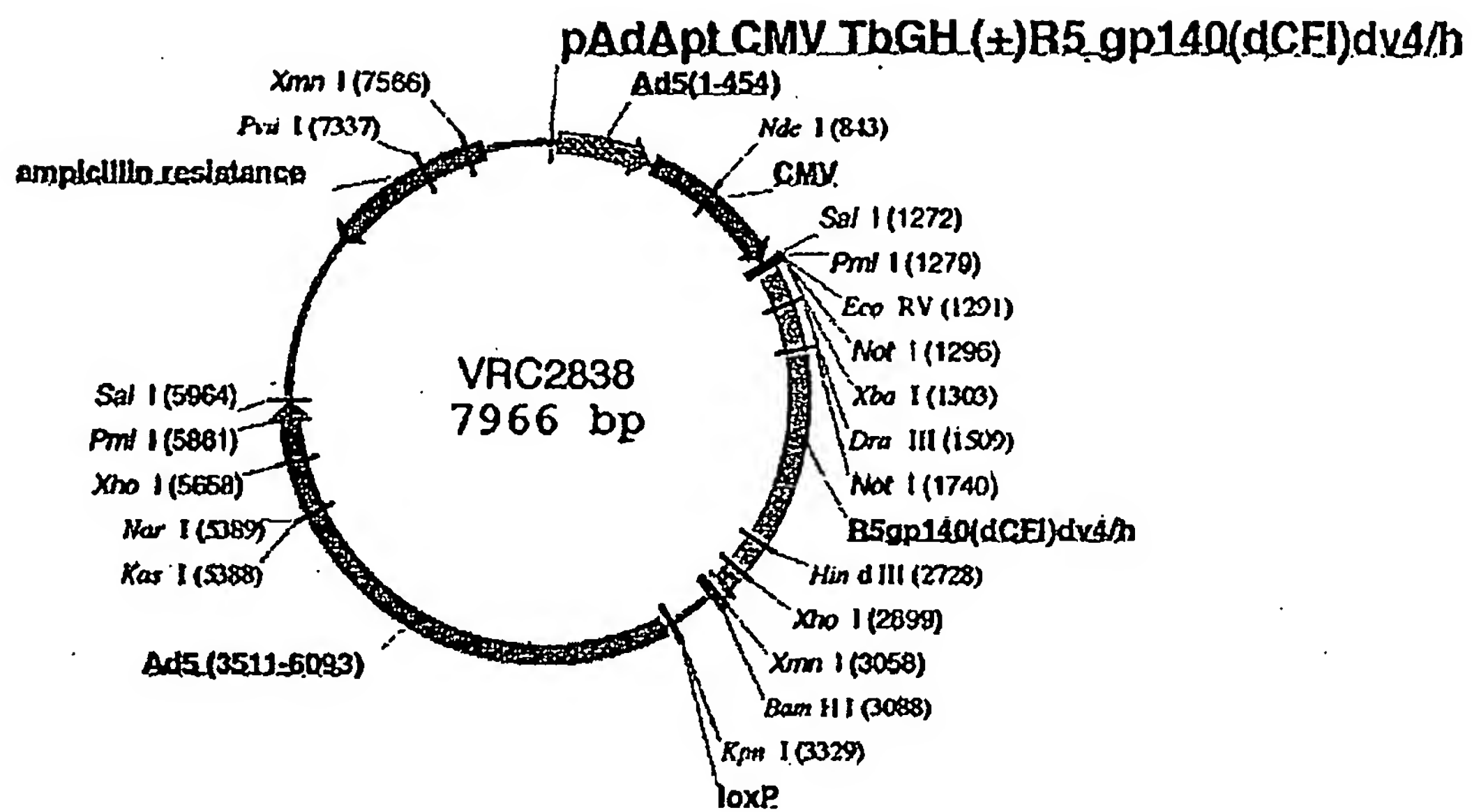


Figure 38

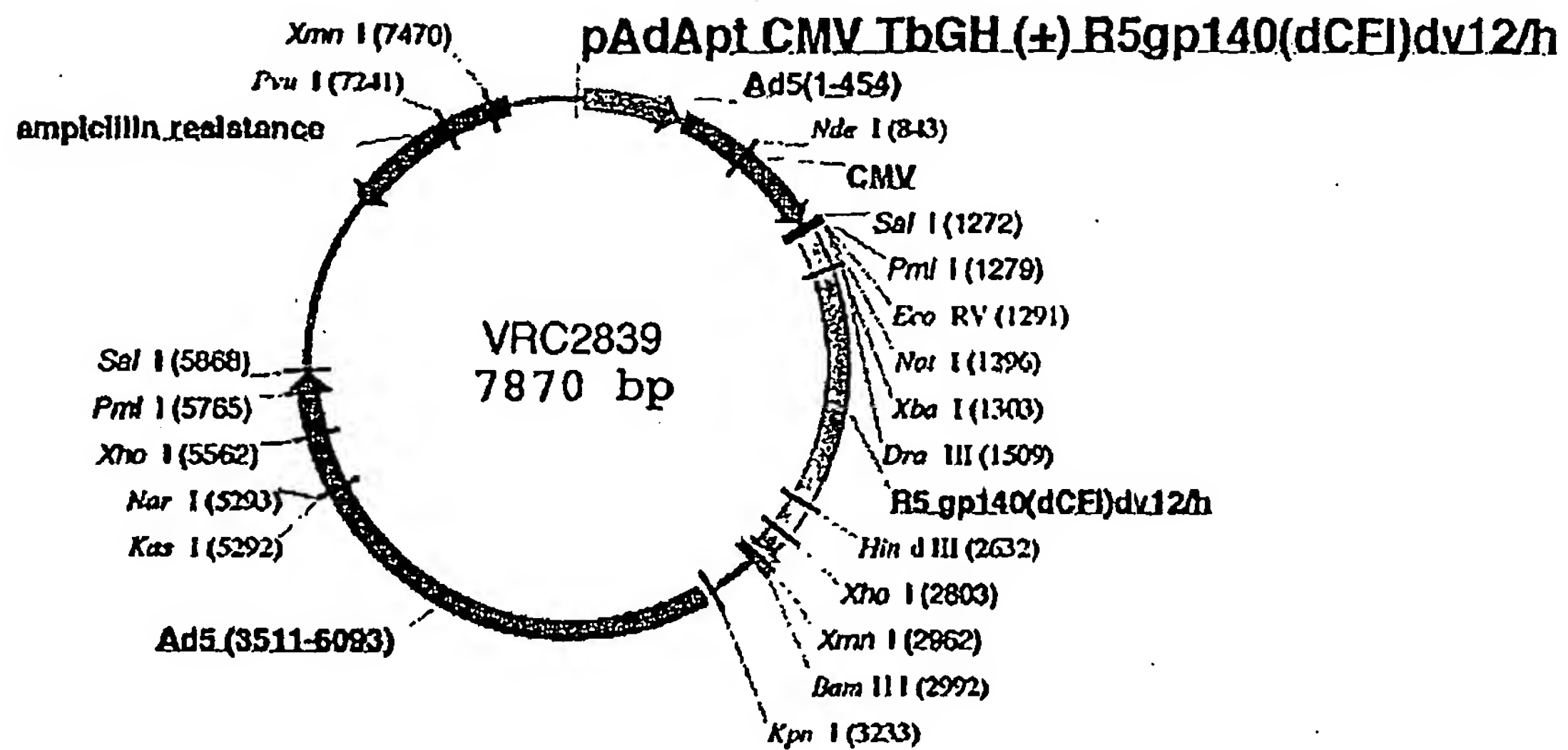


Figure 39

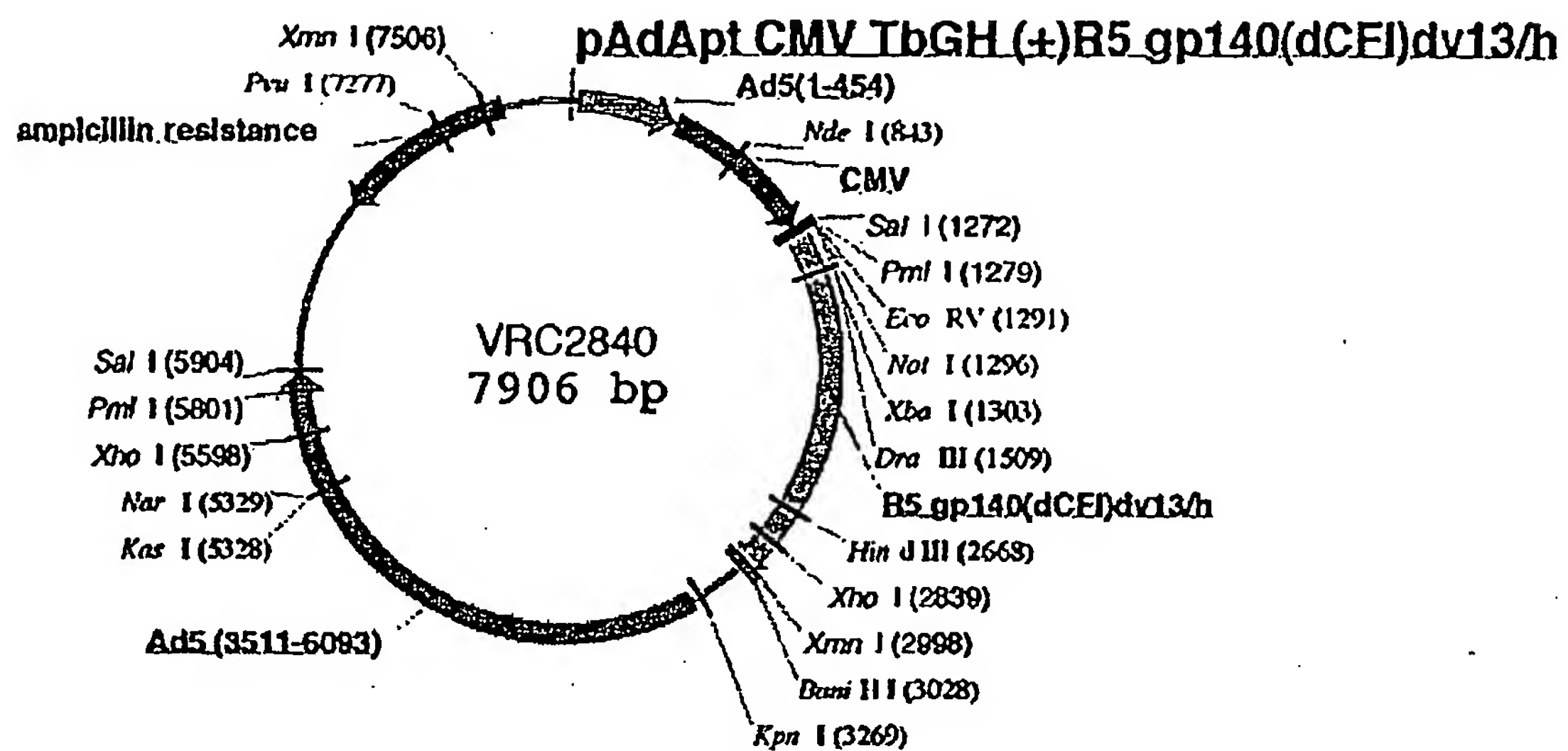


Figure 40

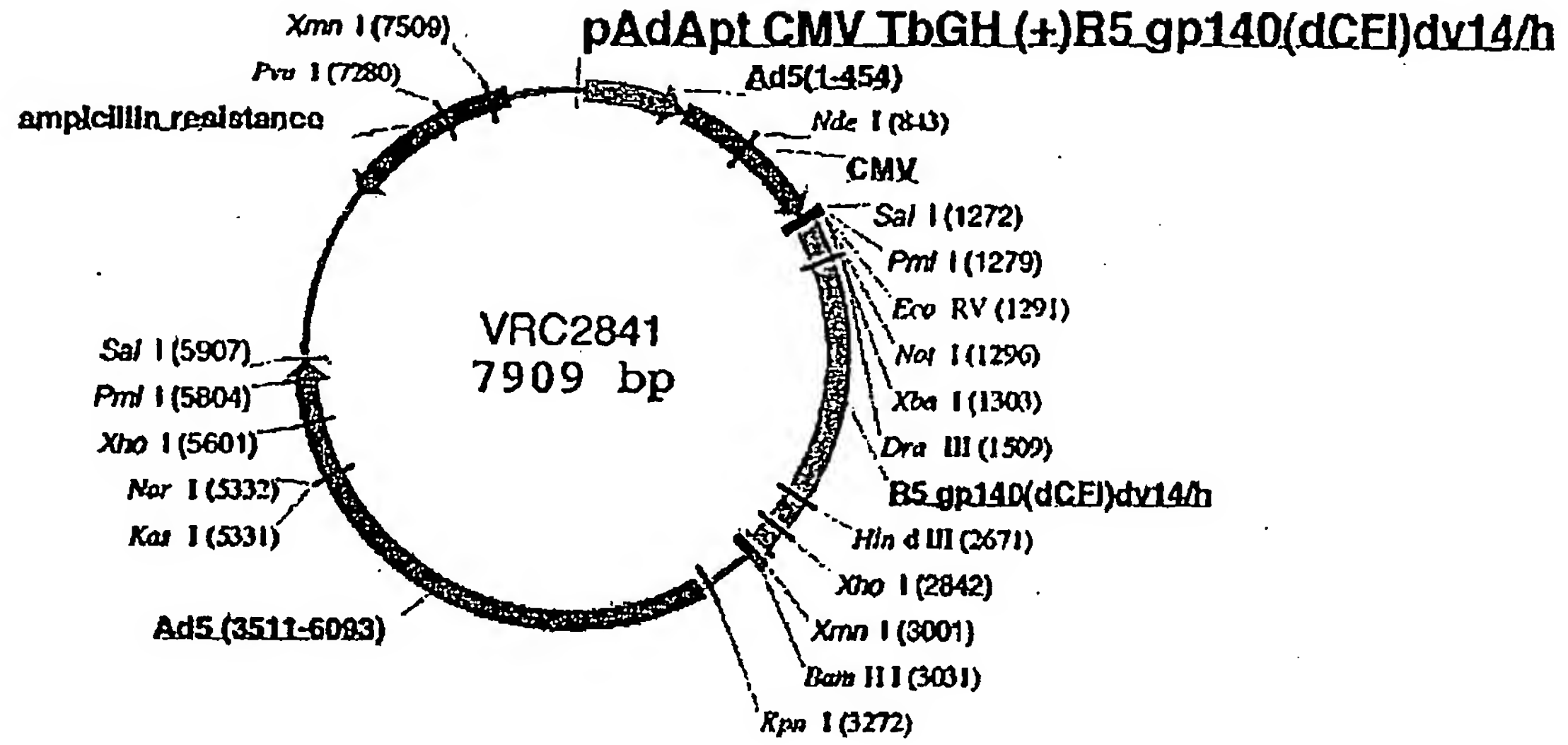


Figure 41



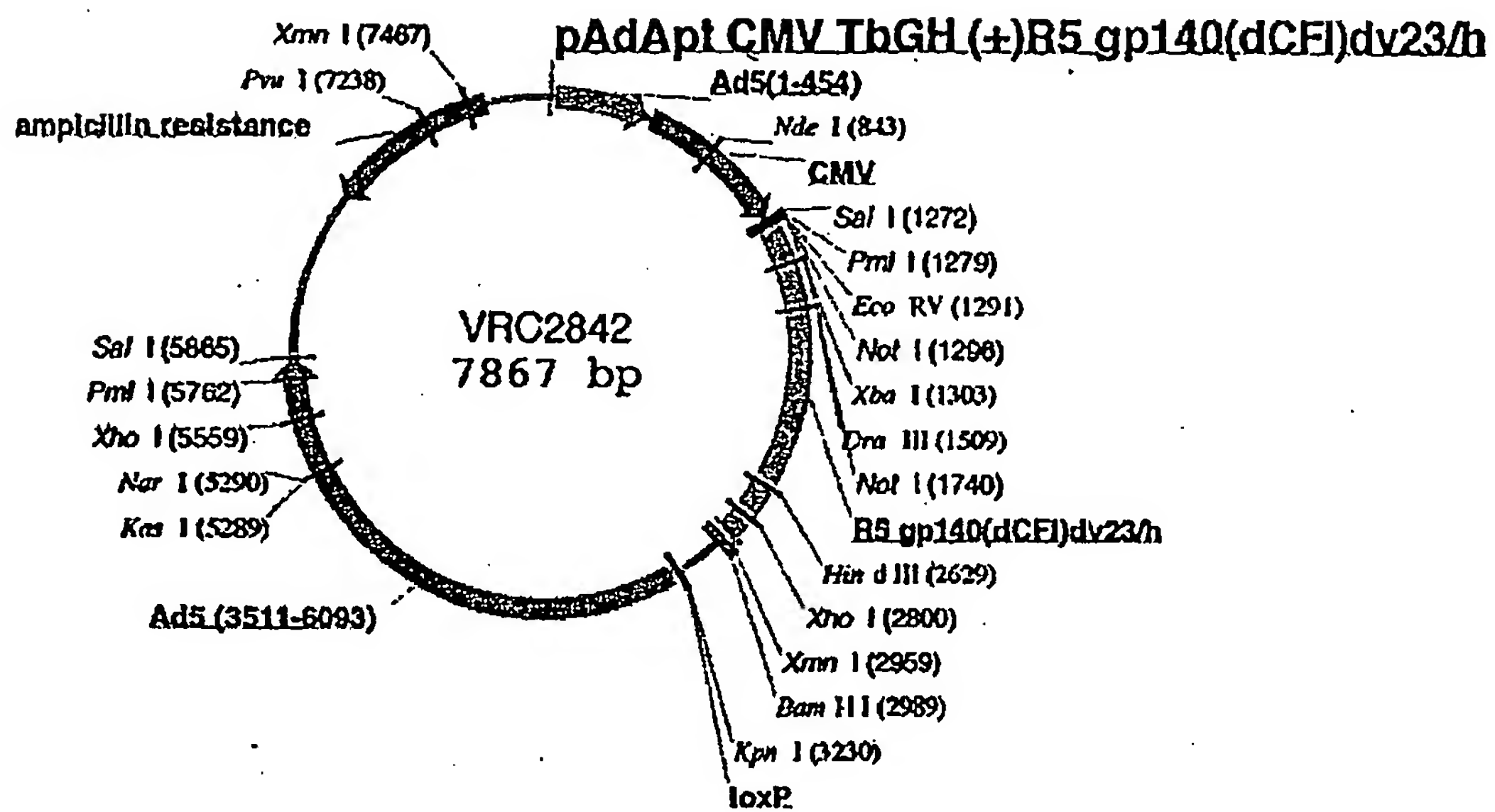


Figure 42

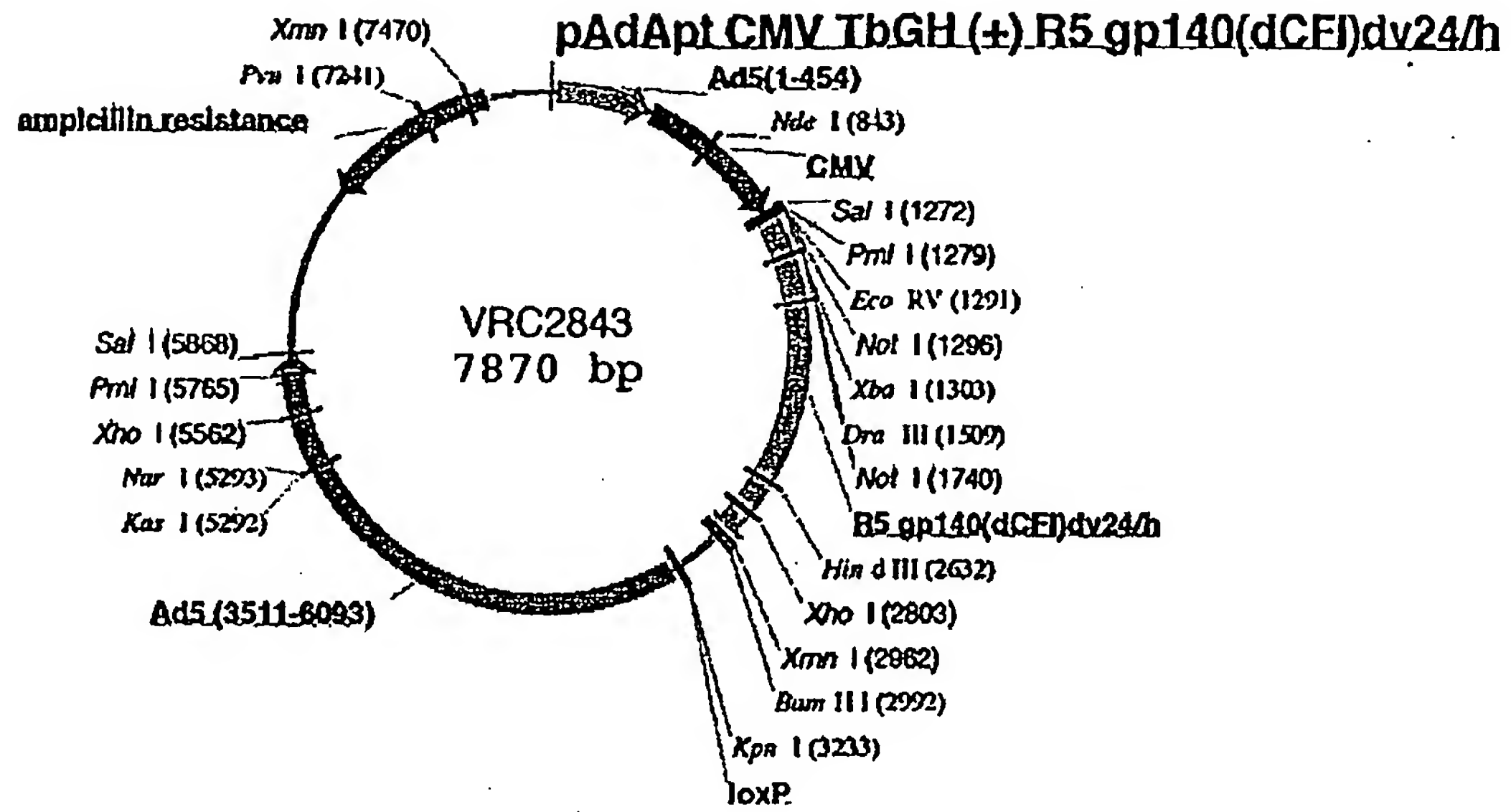


Figure 43

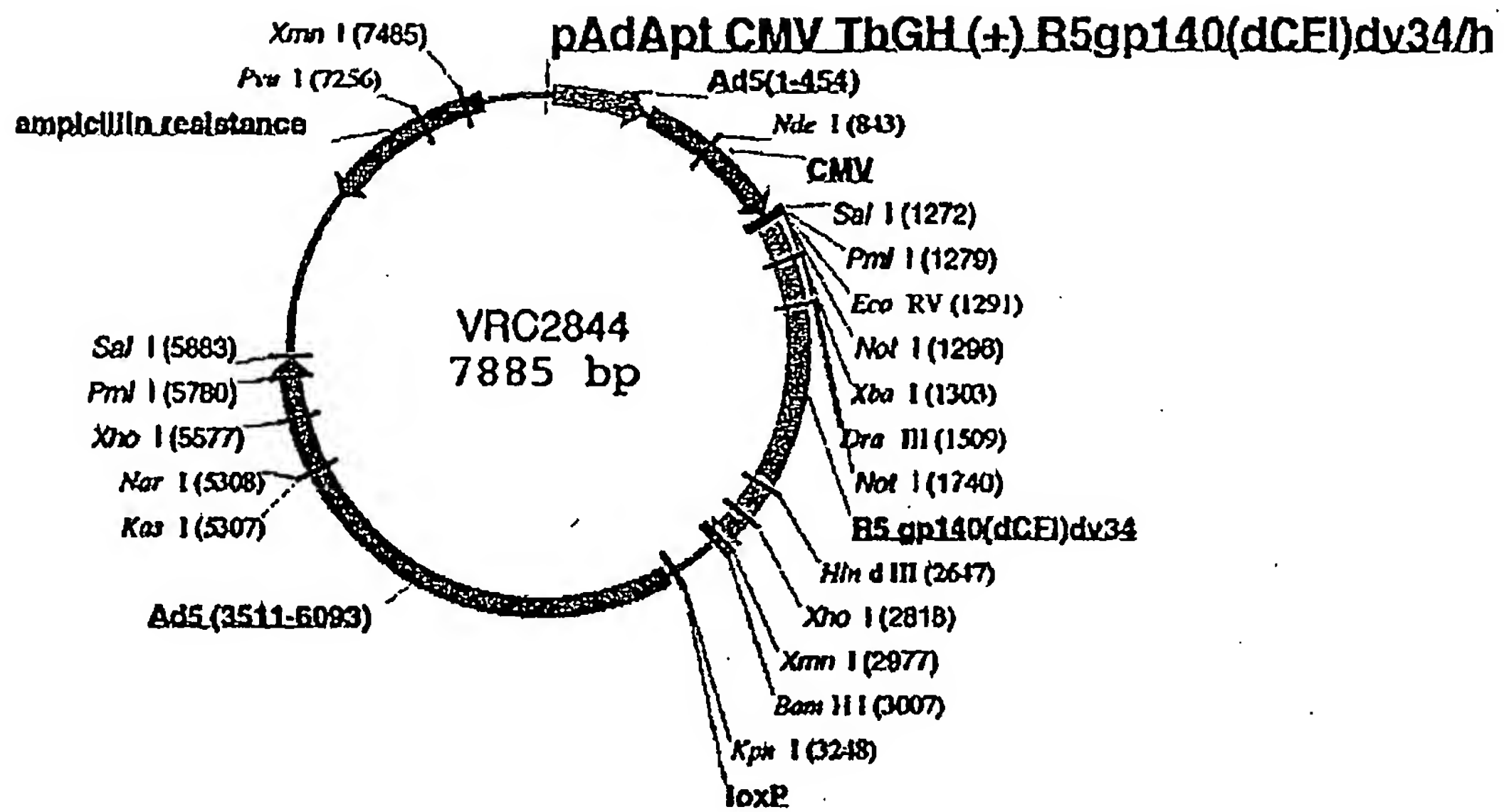


Figure 44

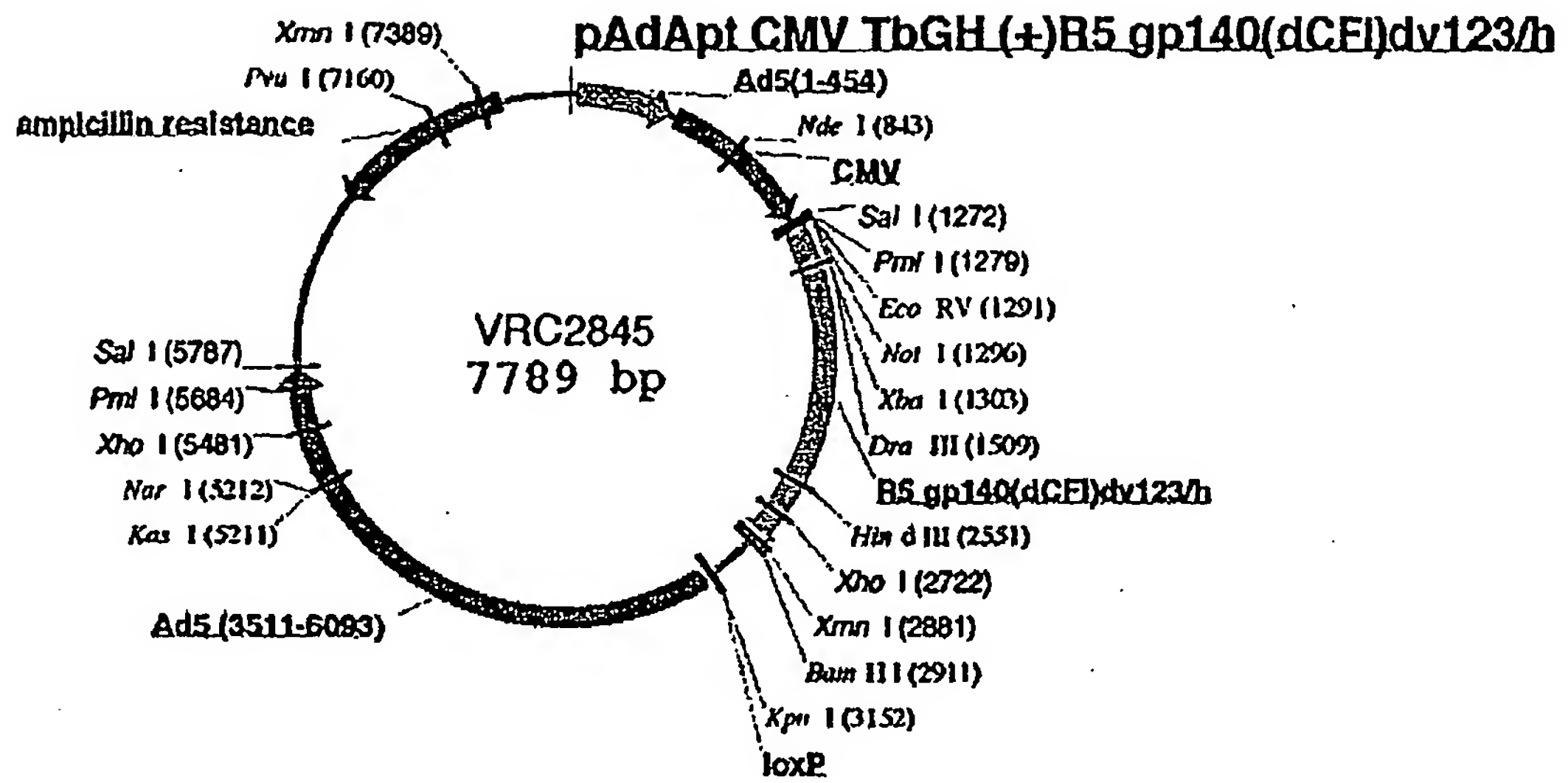


Figure 45

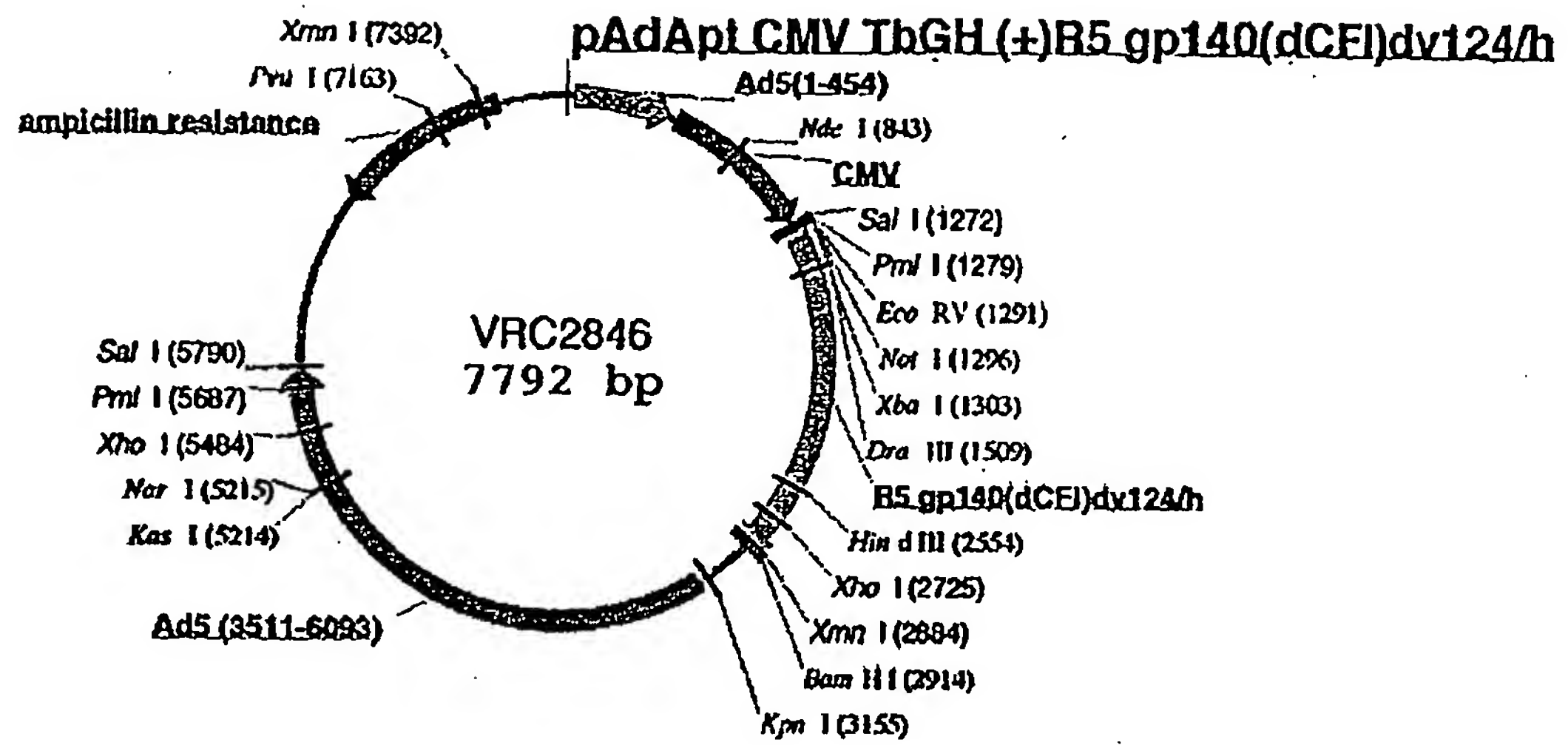


Figure 46

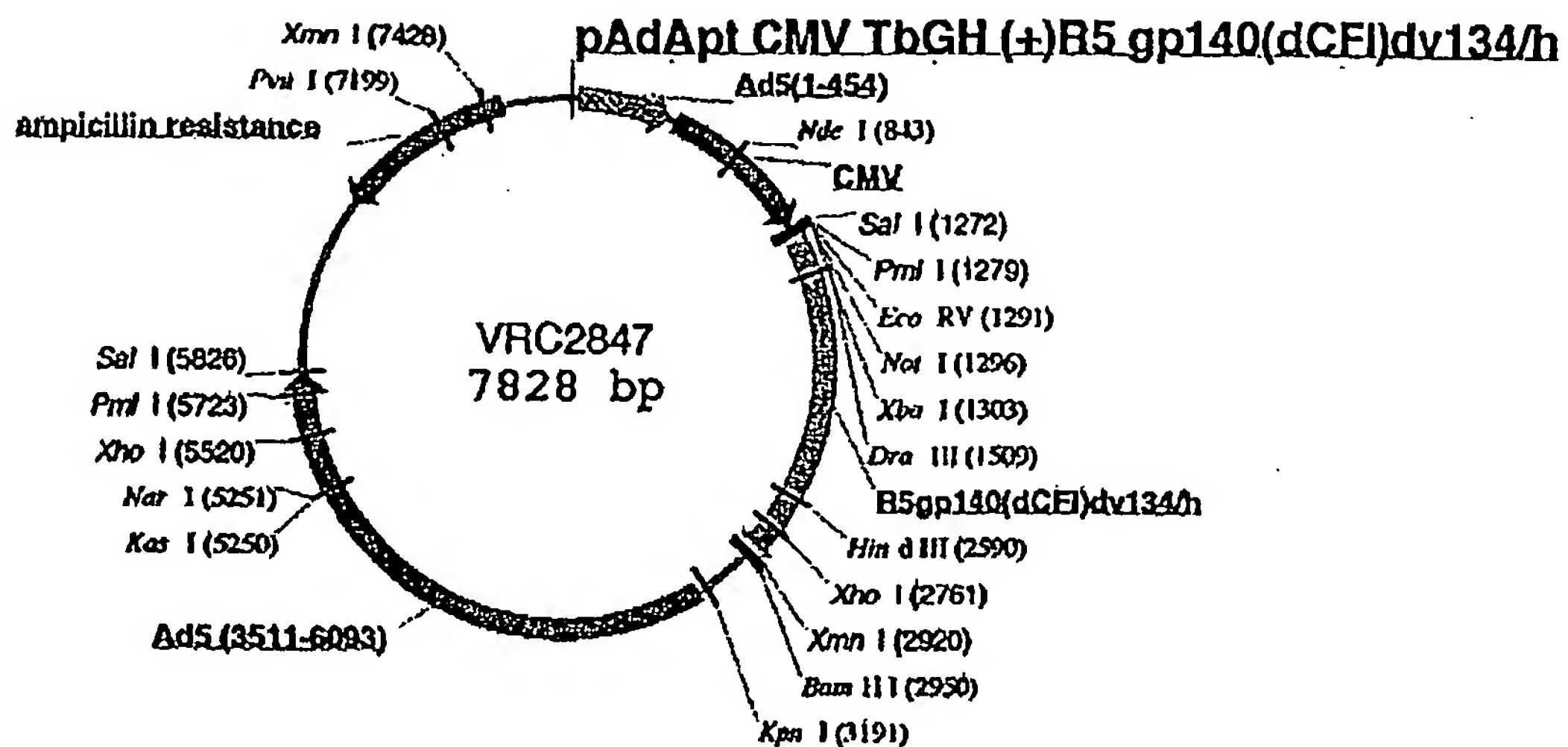


Figure 47



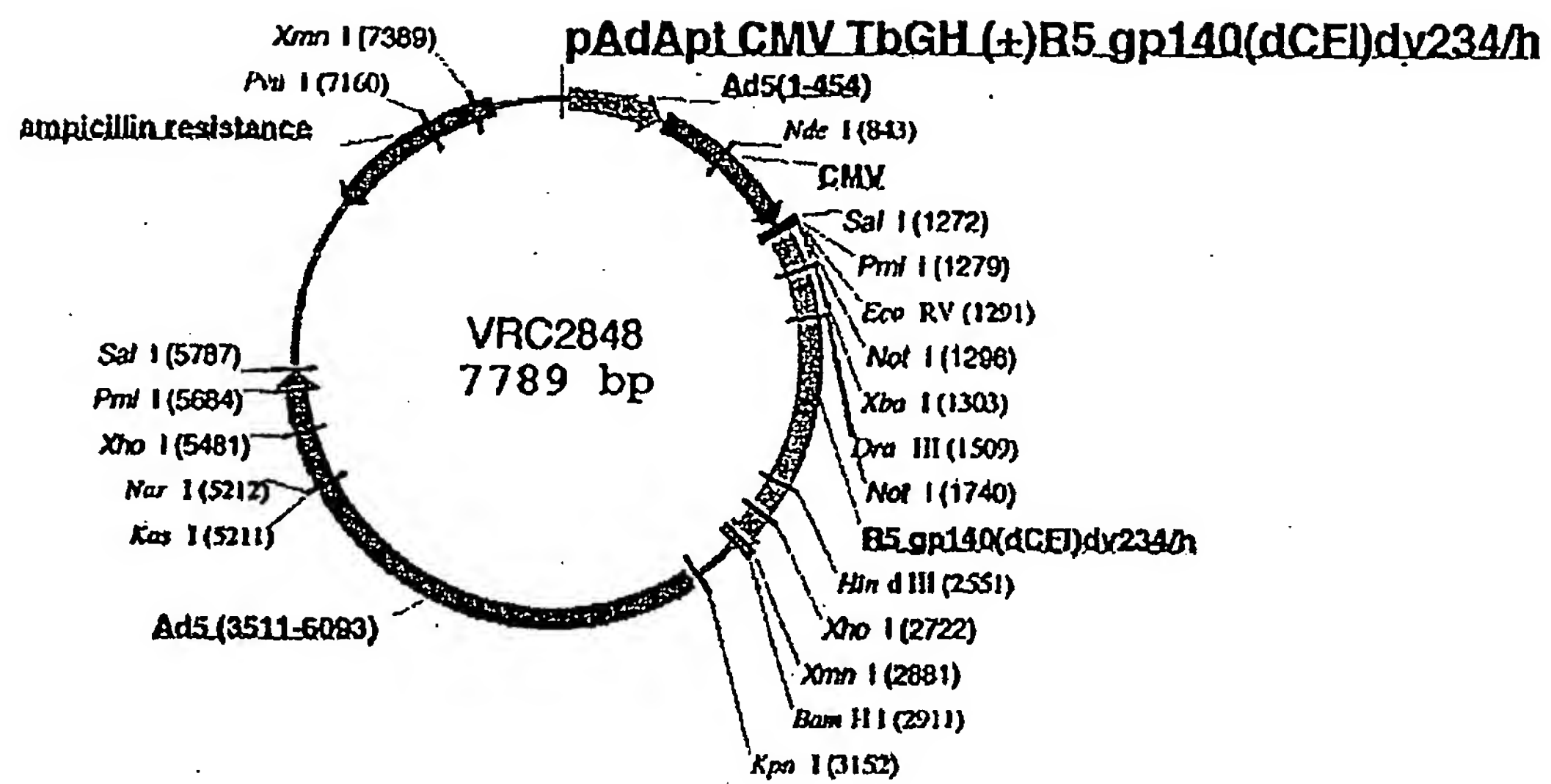


Figure 48

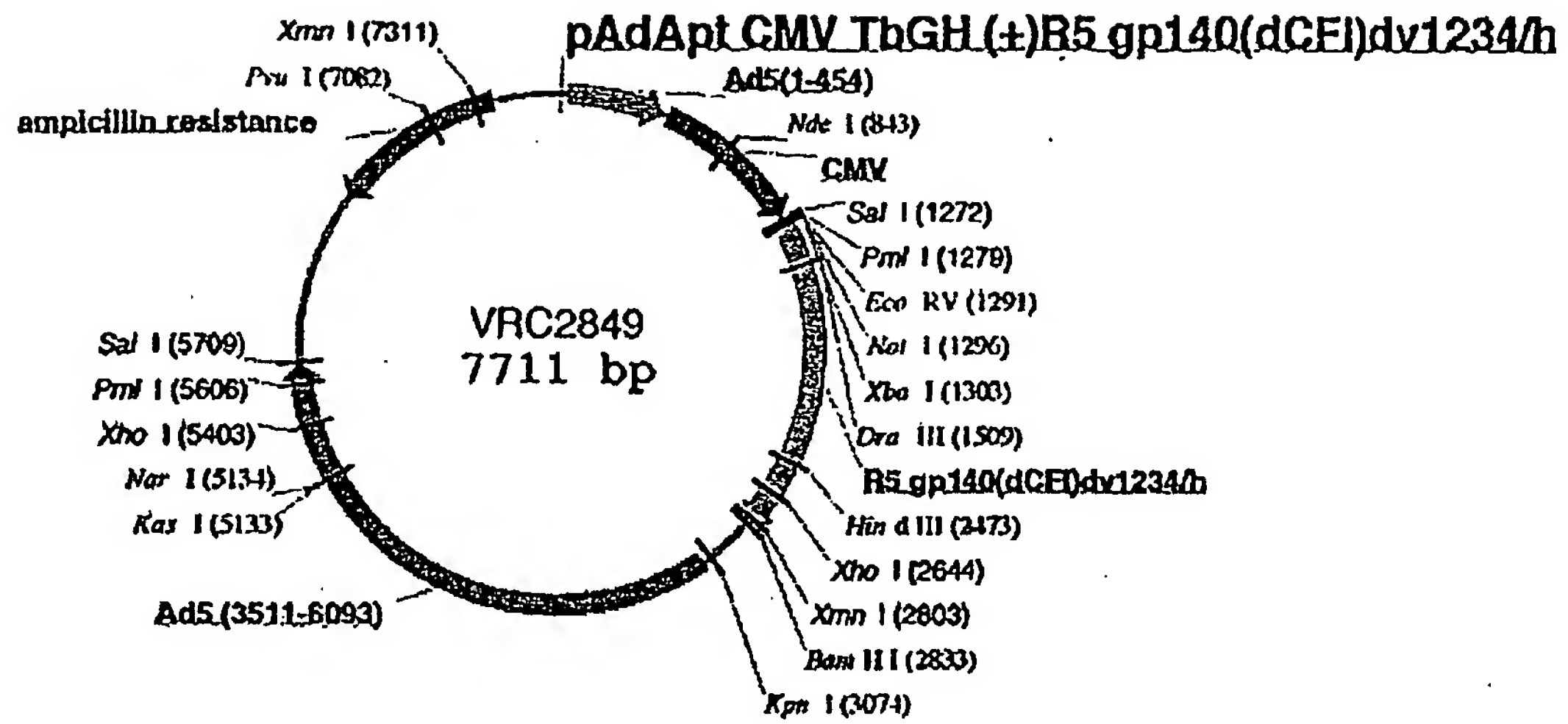


Figure 49

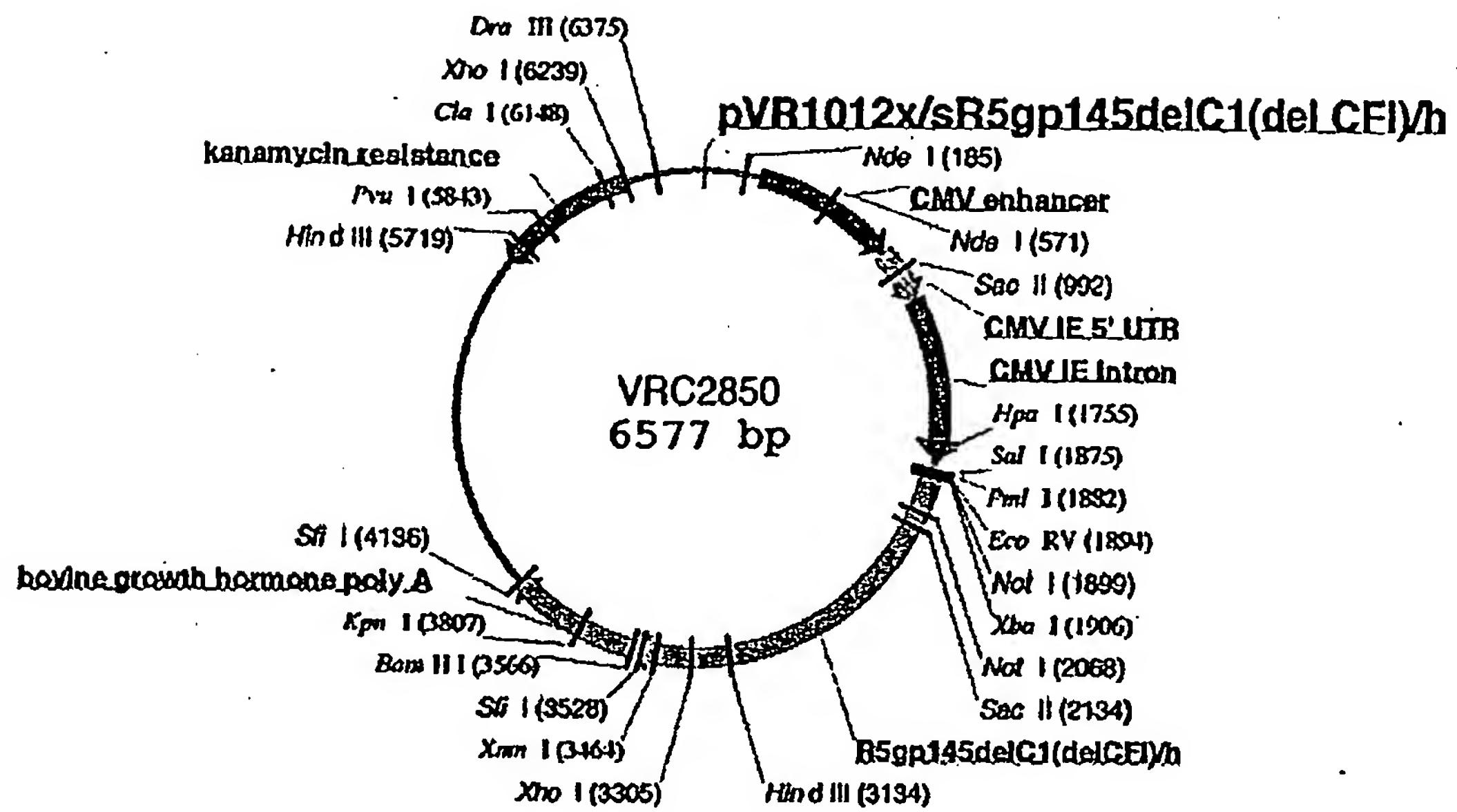


Figure 50

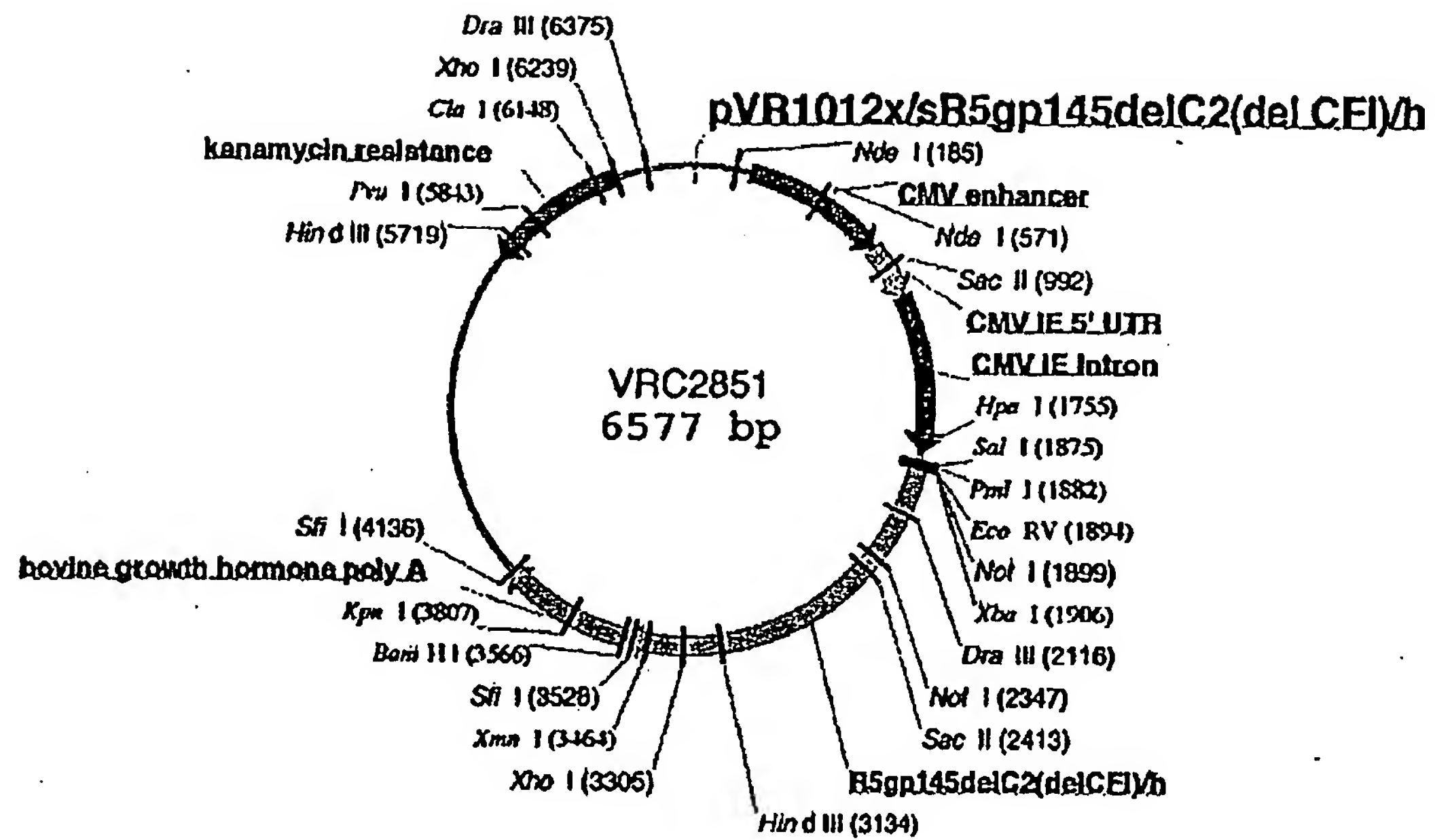


Figure 51

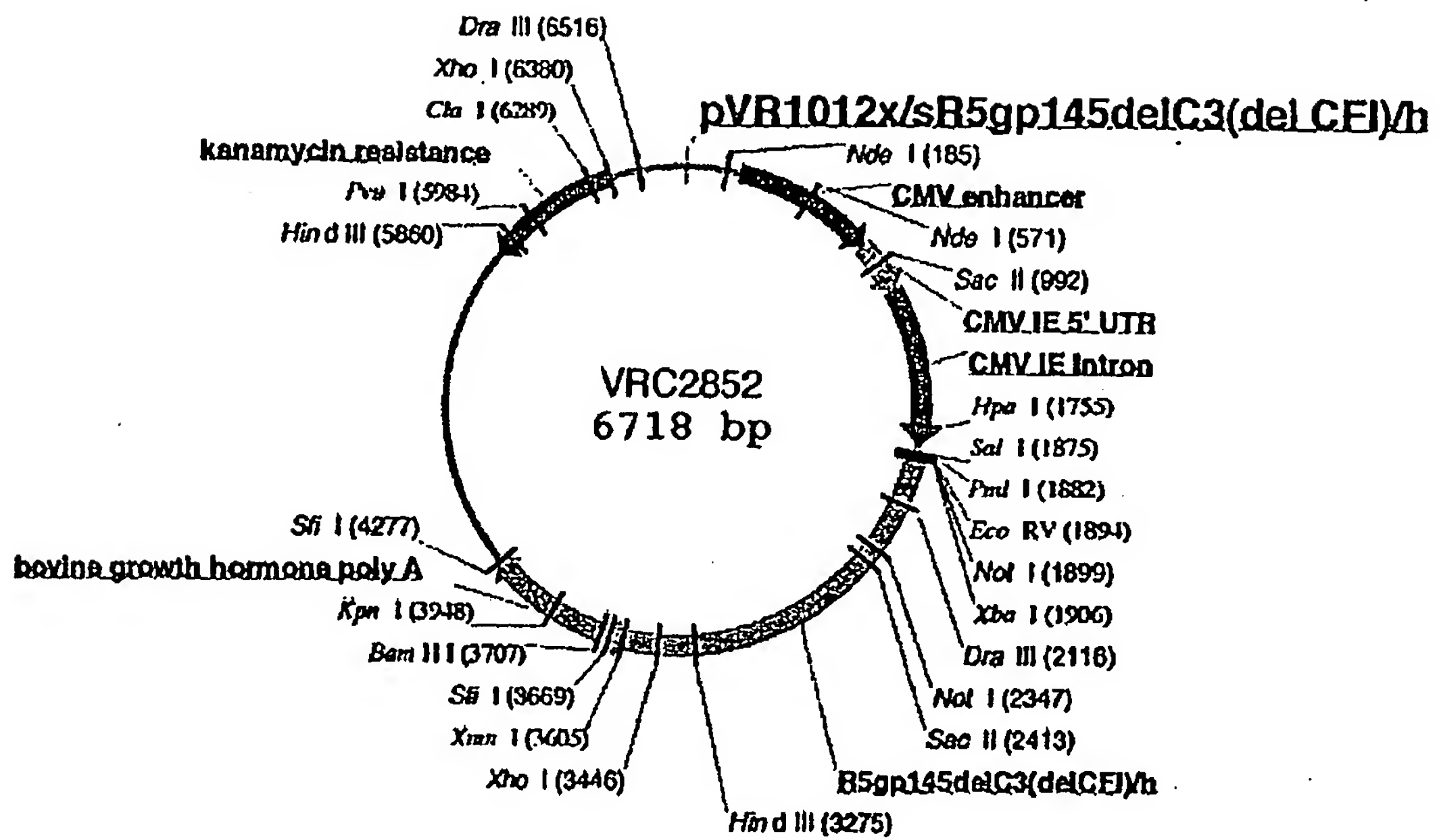


Figure 52

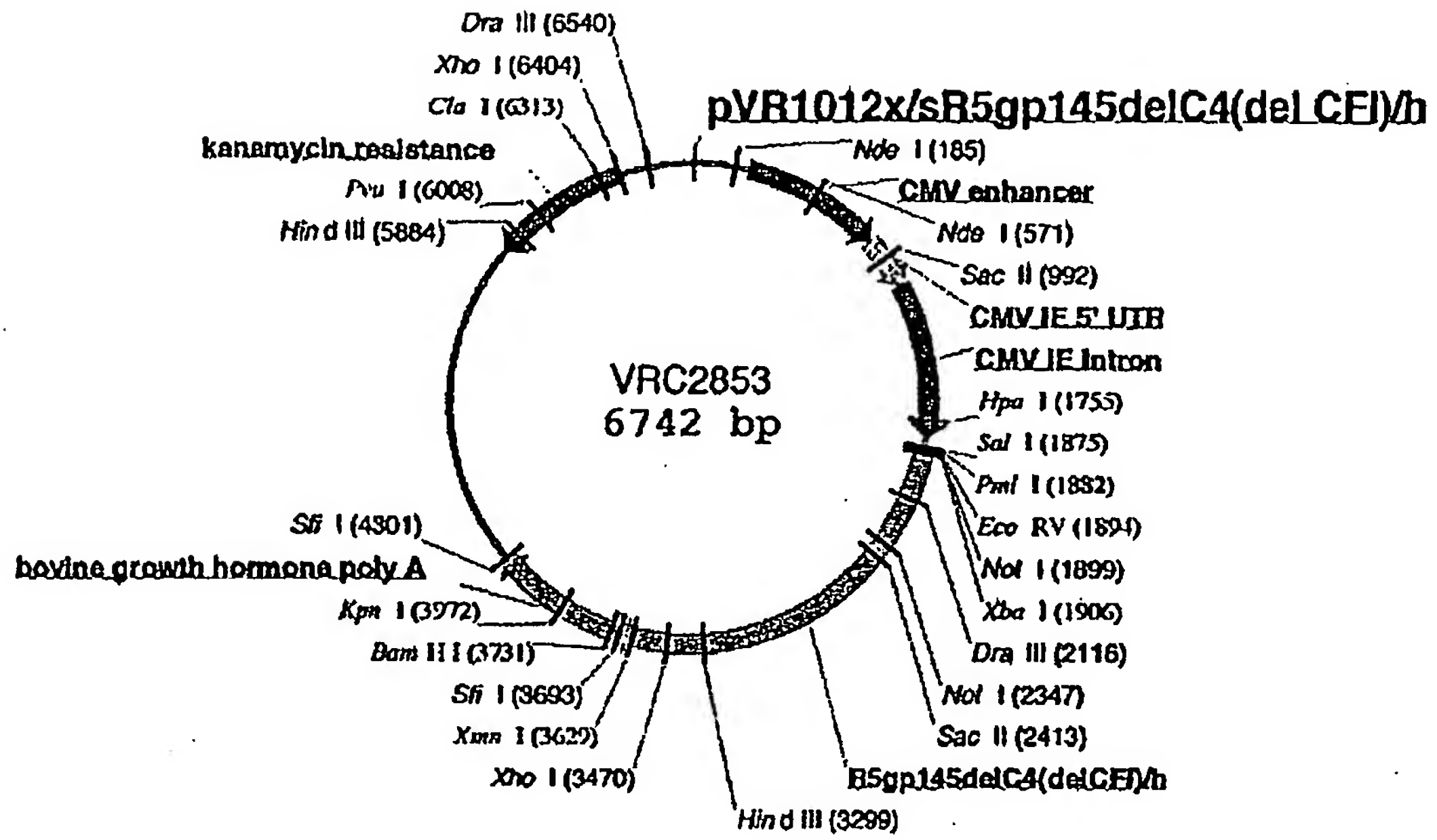


Figure 53



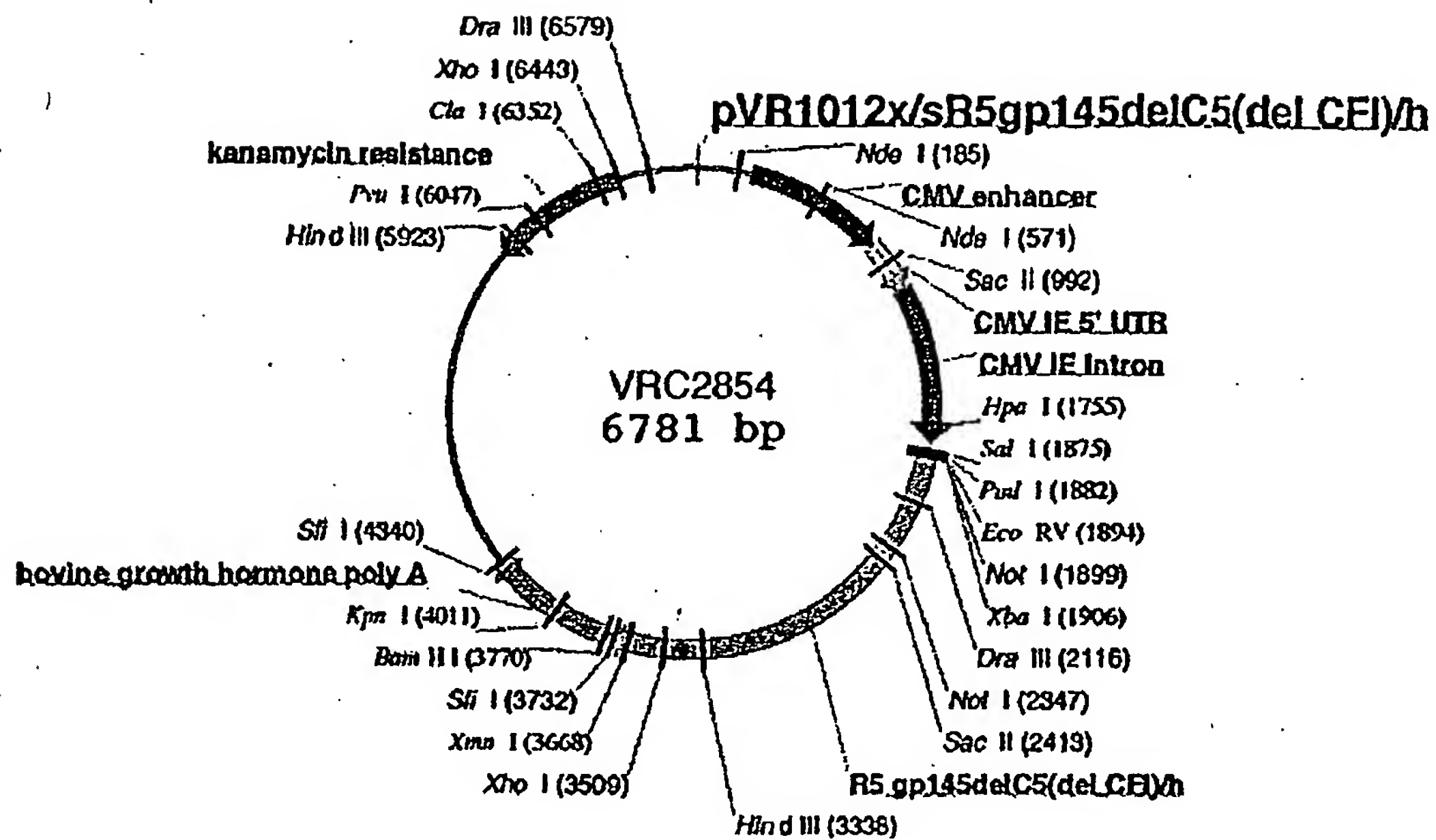


Figure 54

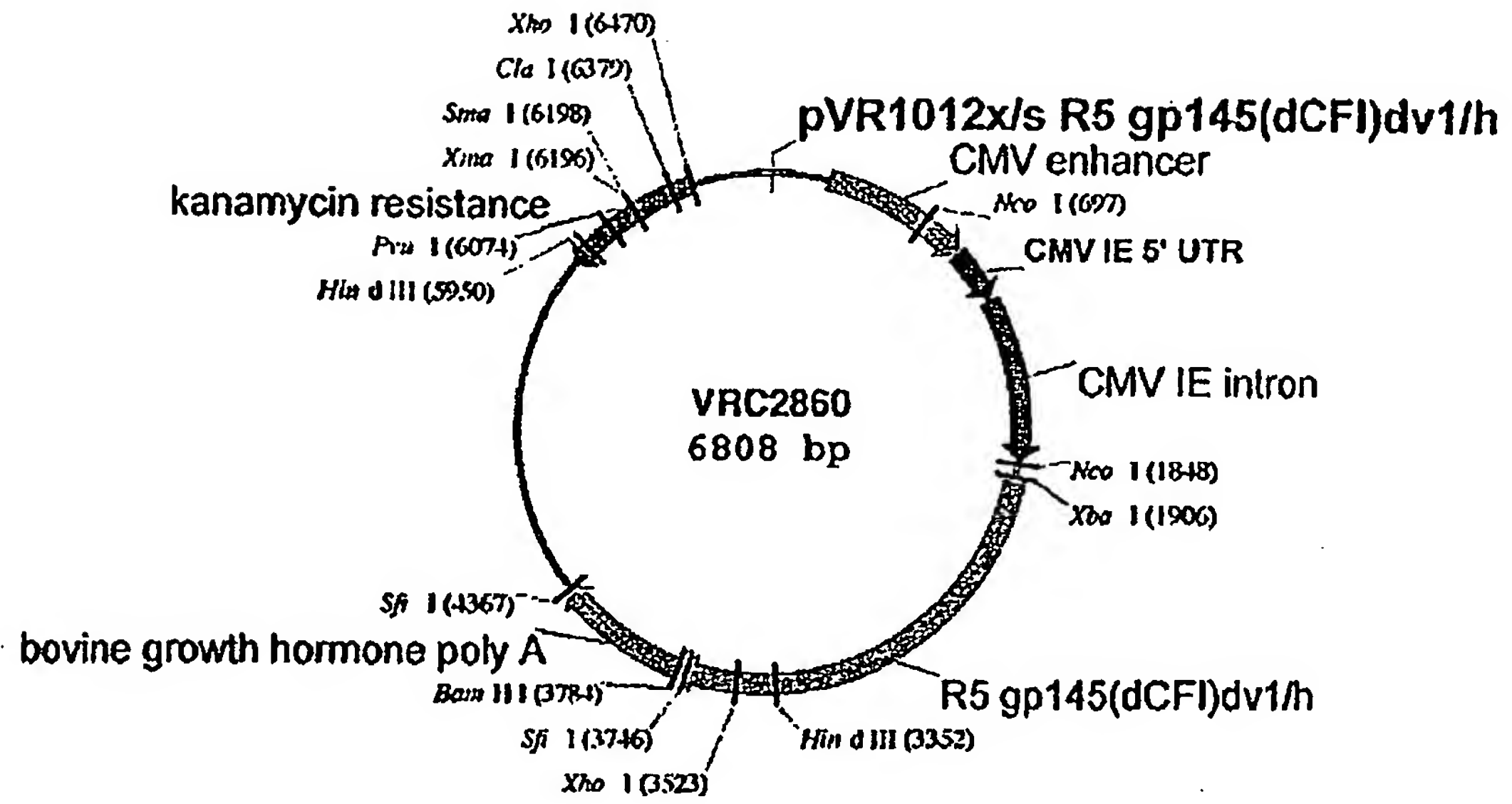


Figure 55

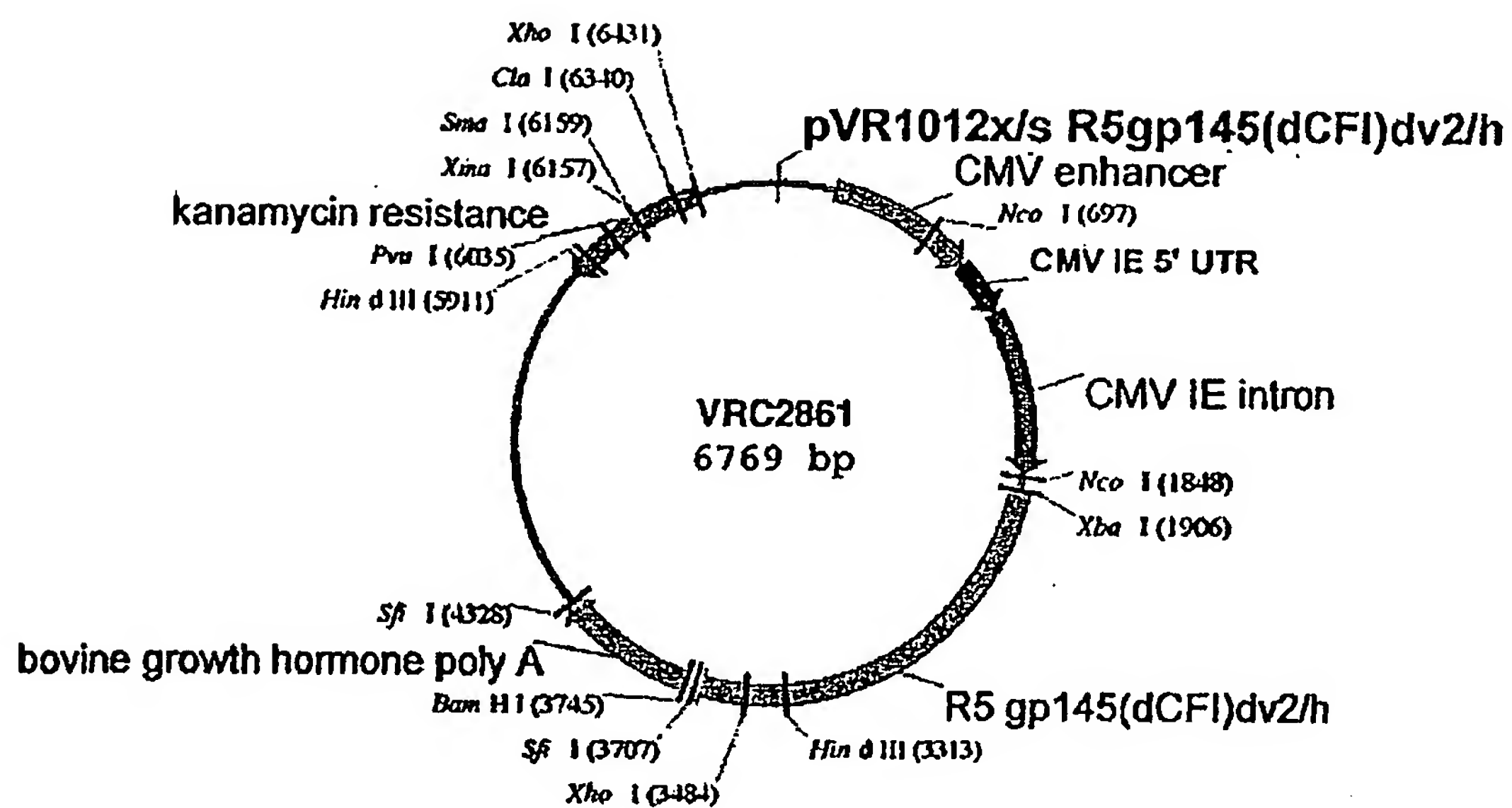


Figure 56

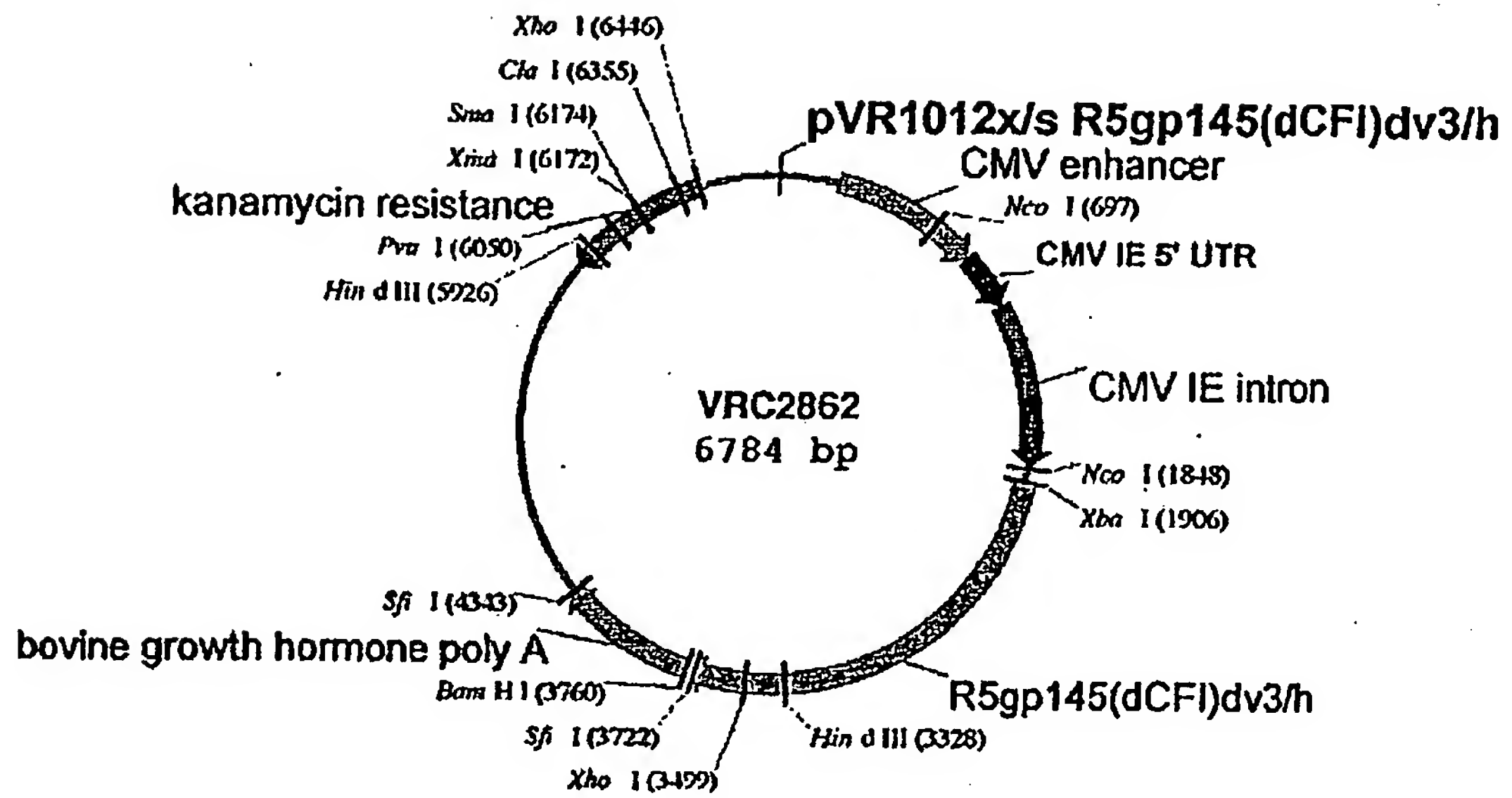


Figure 57

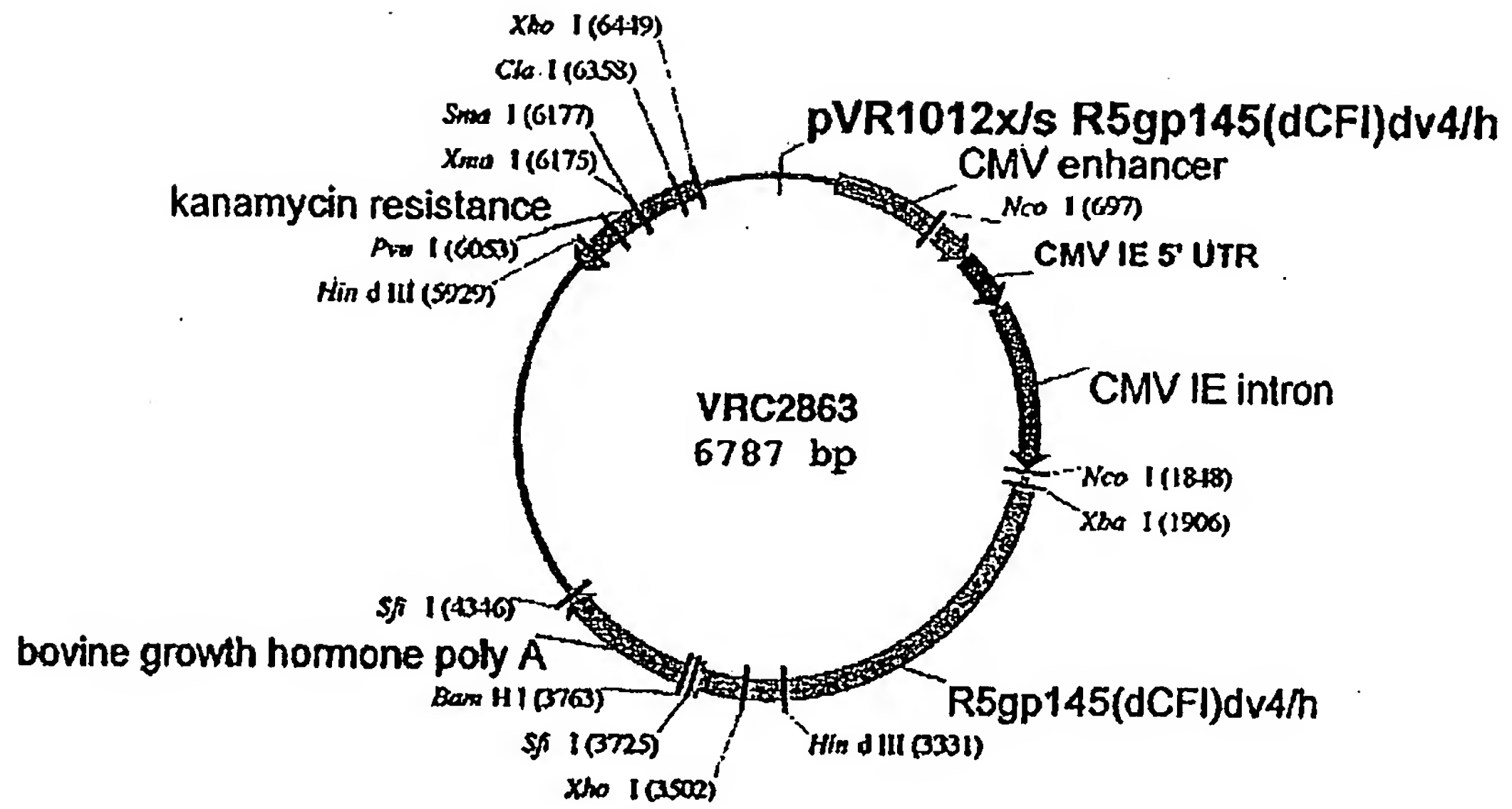


Figure 58

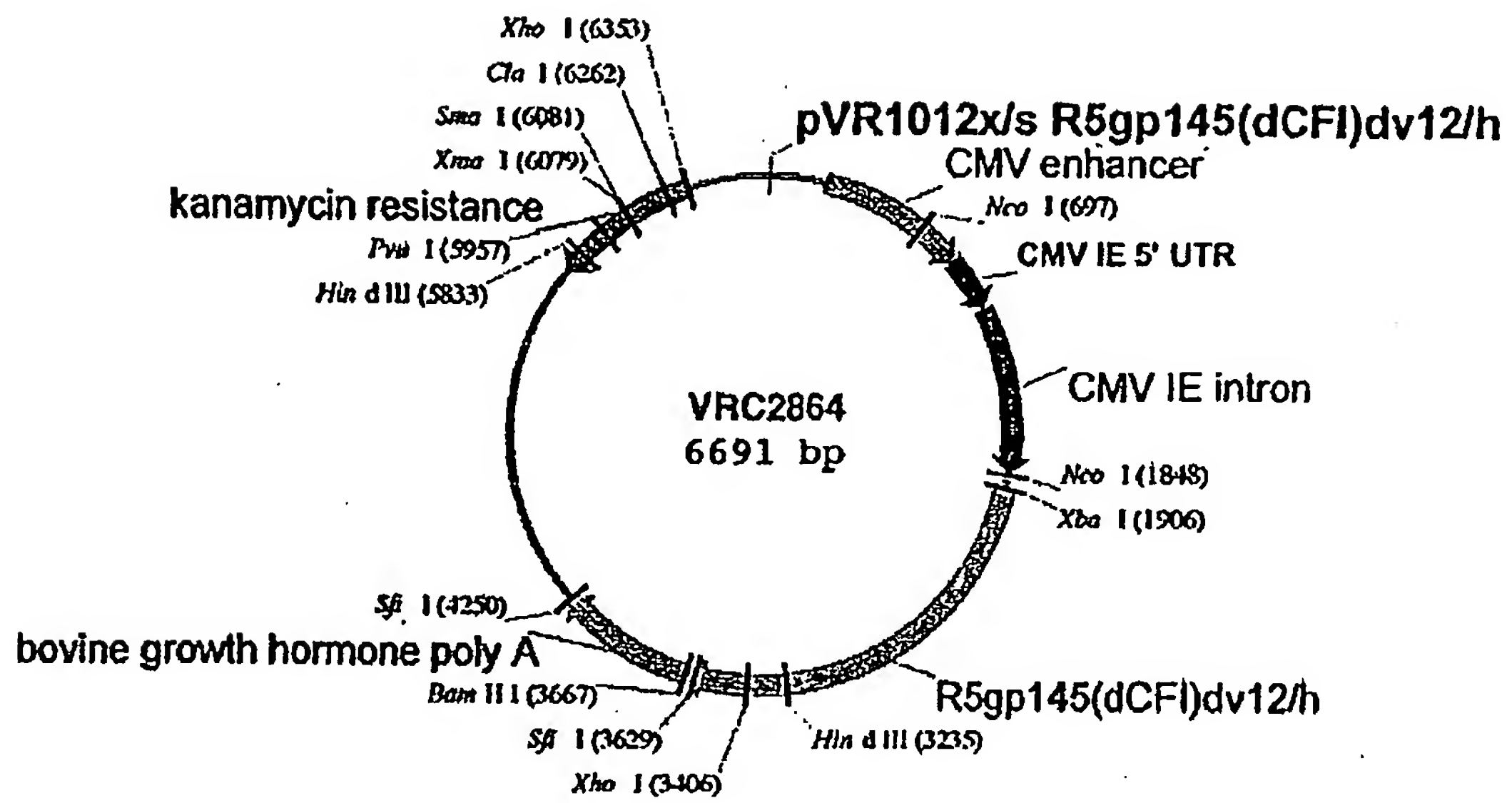


Figure 59



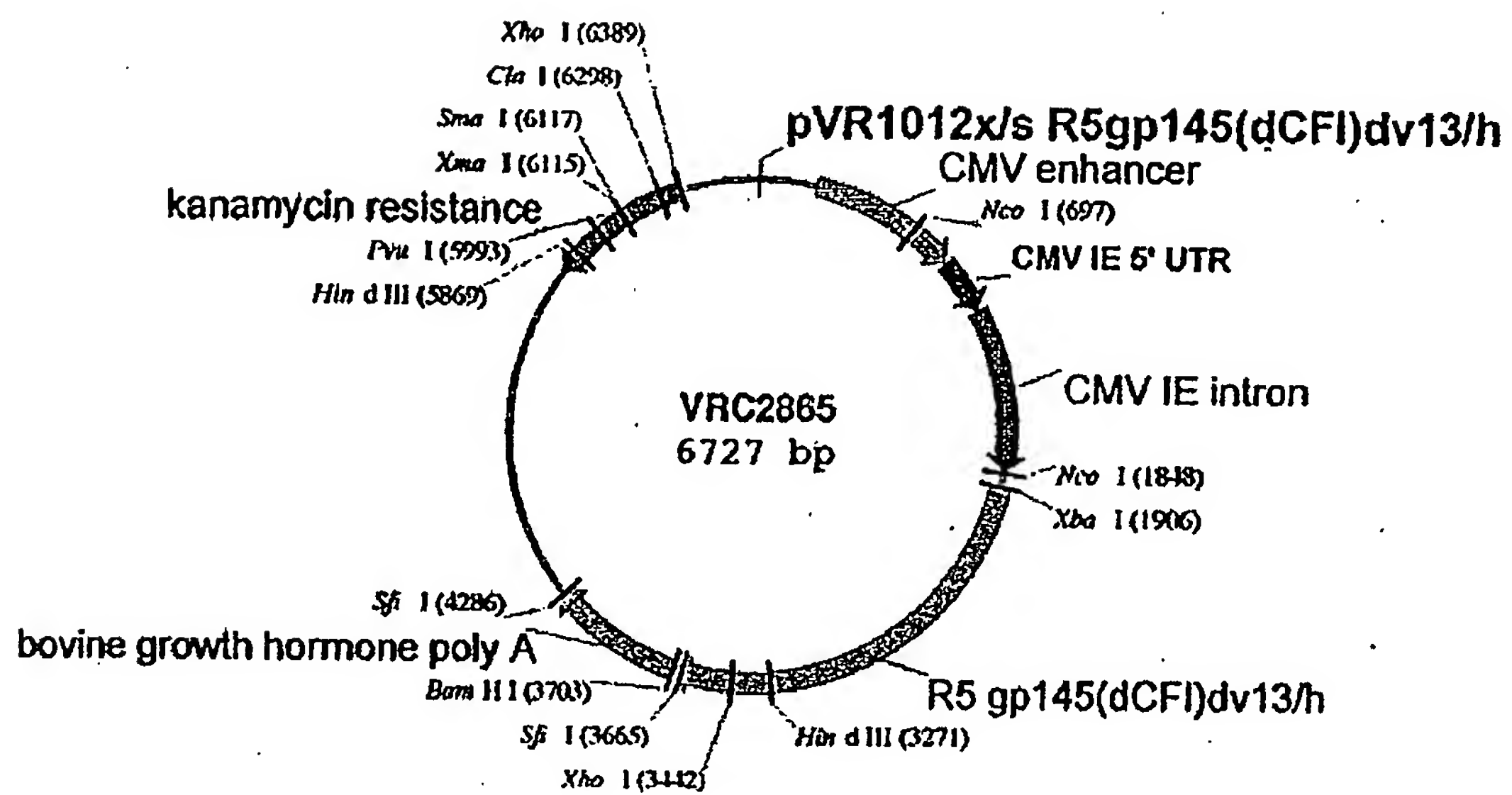


Figure 60

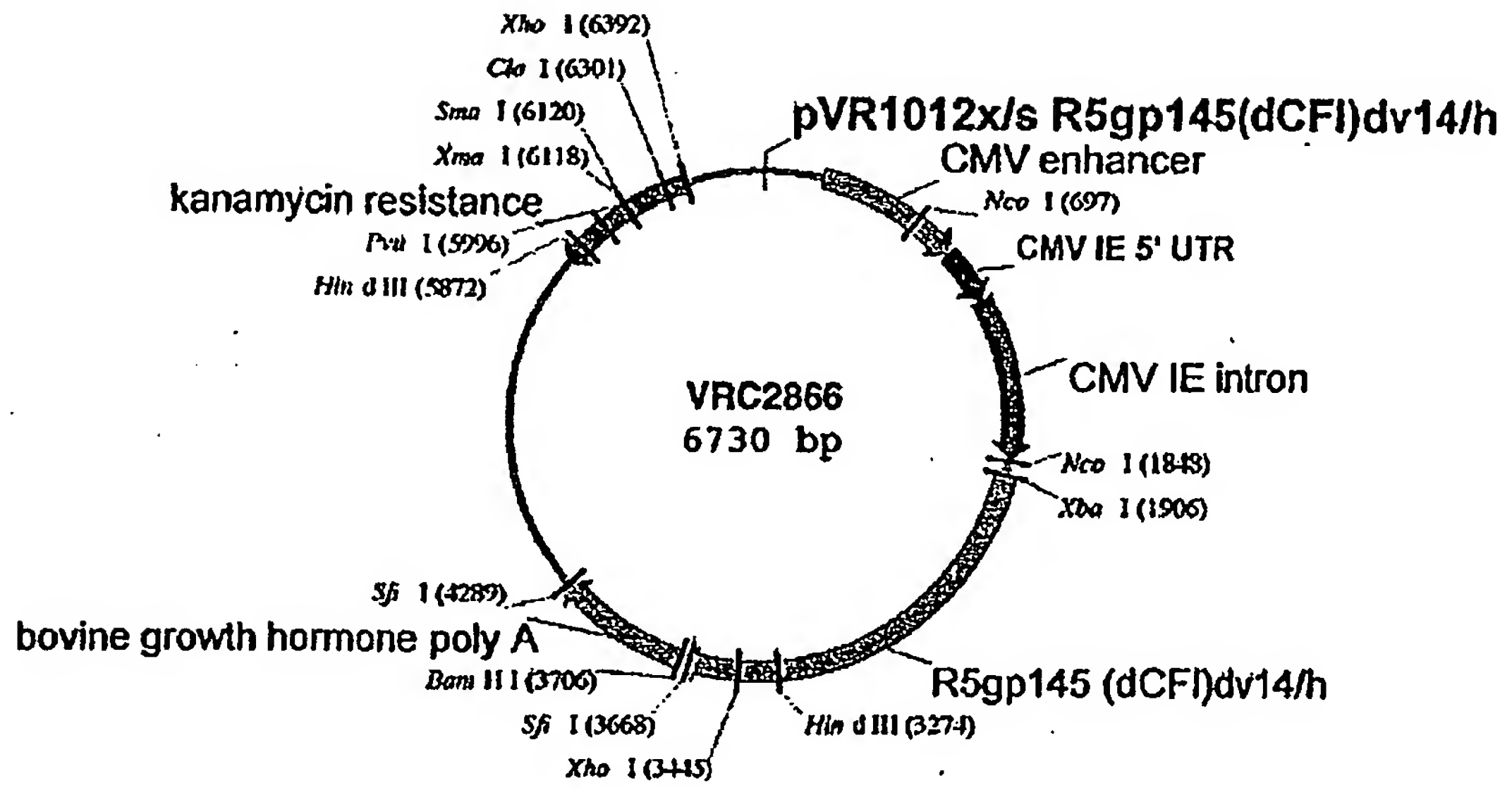


Figure 61

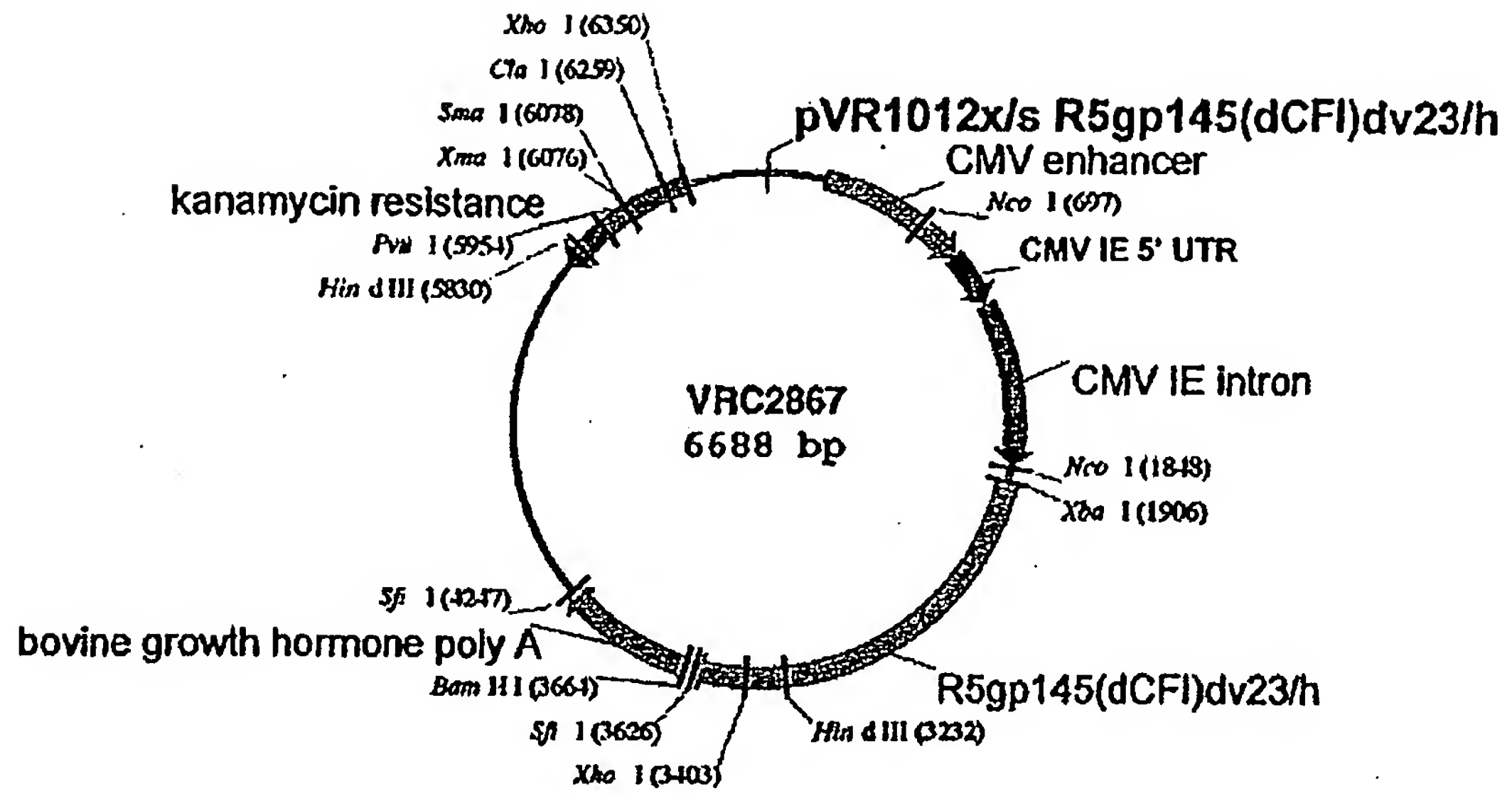


Figure 62

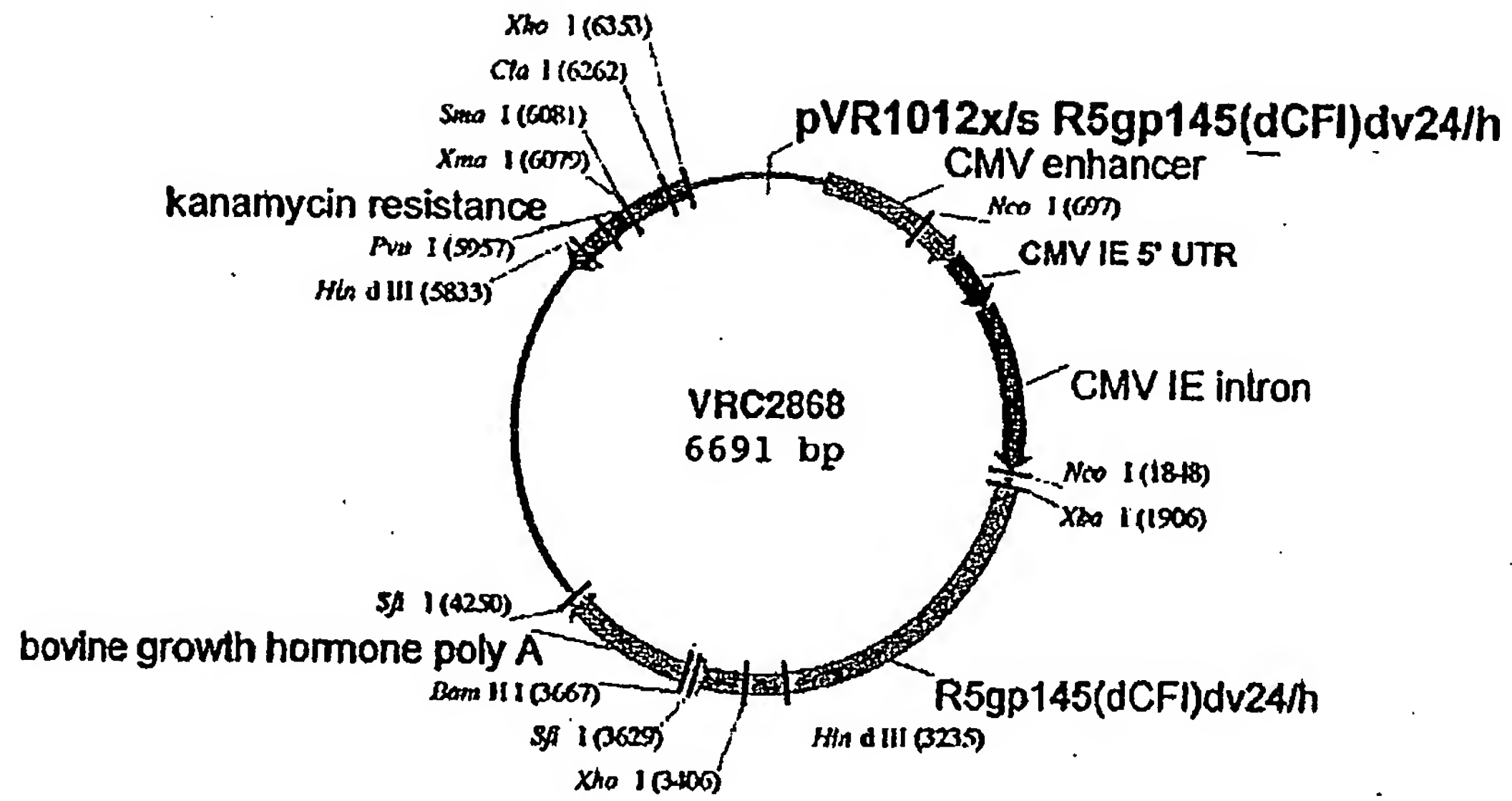


Figure 63

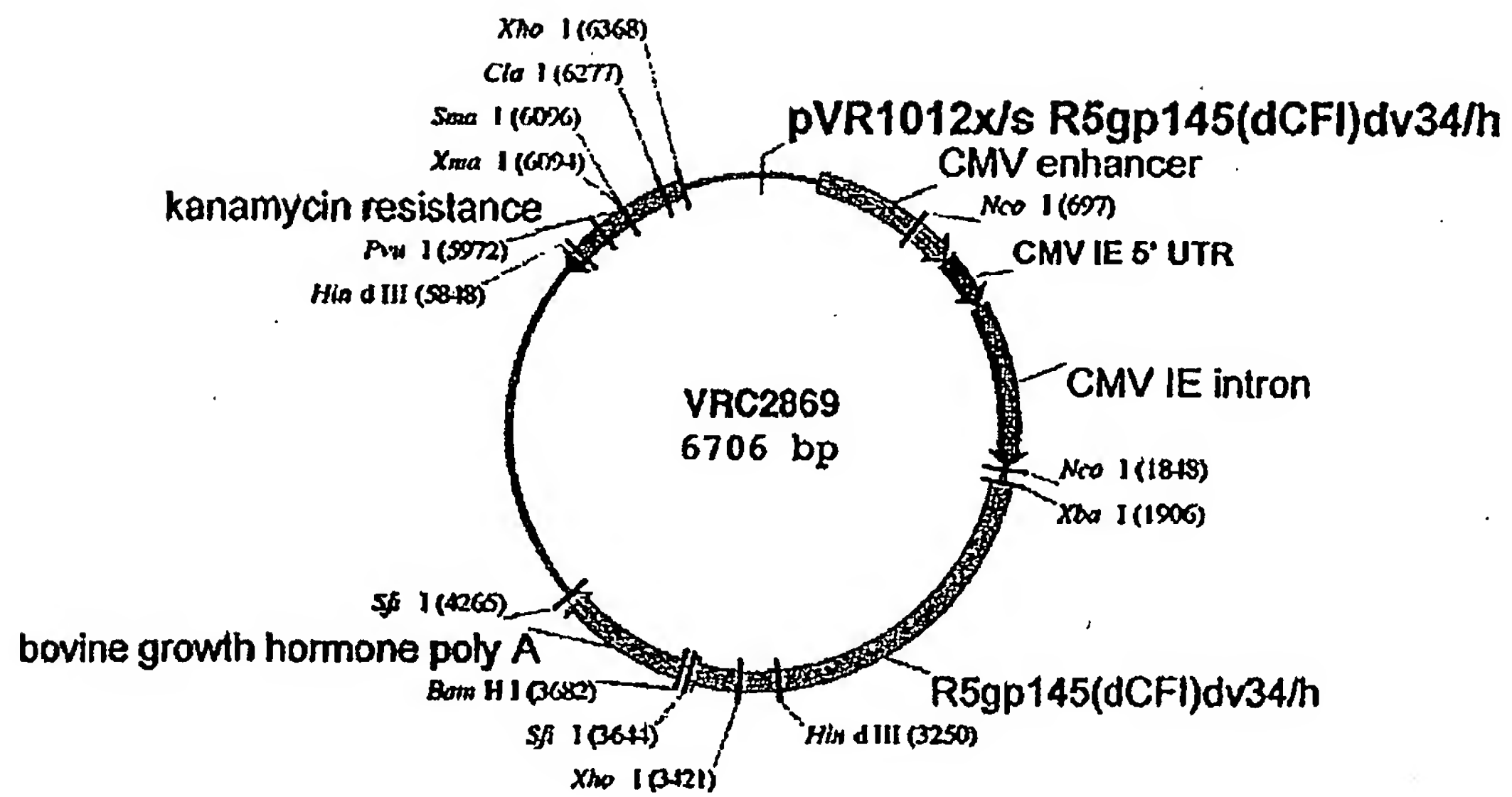


Figure 64

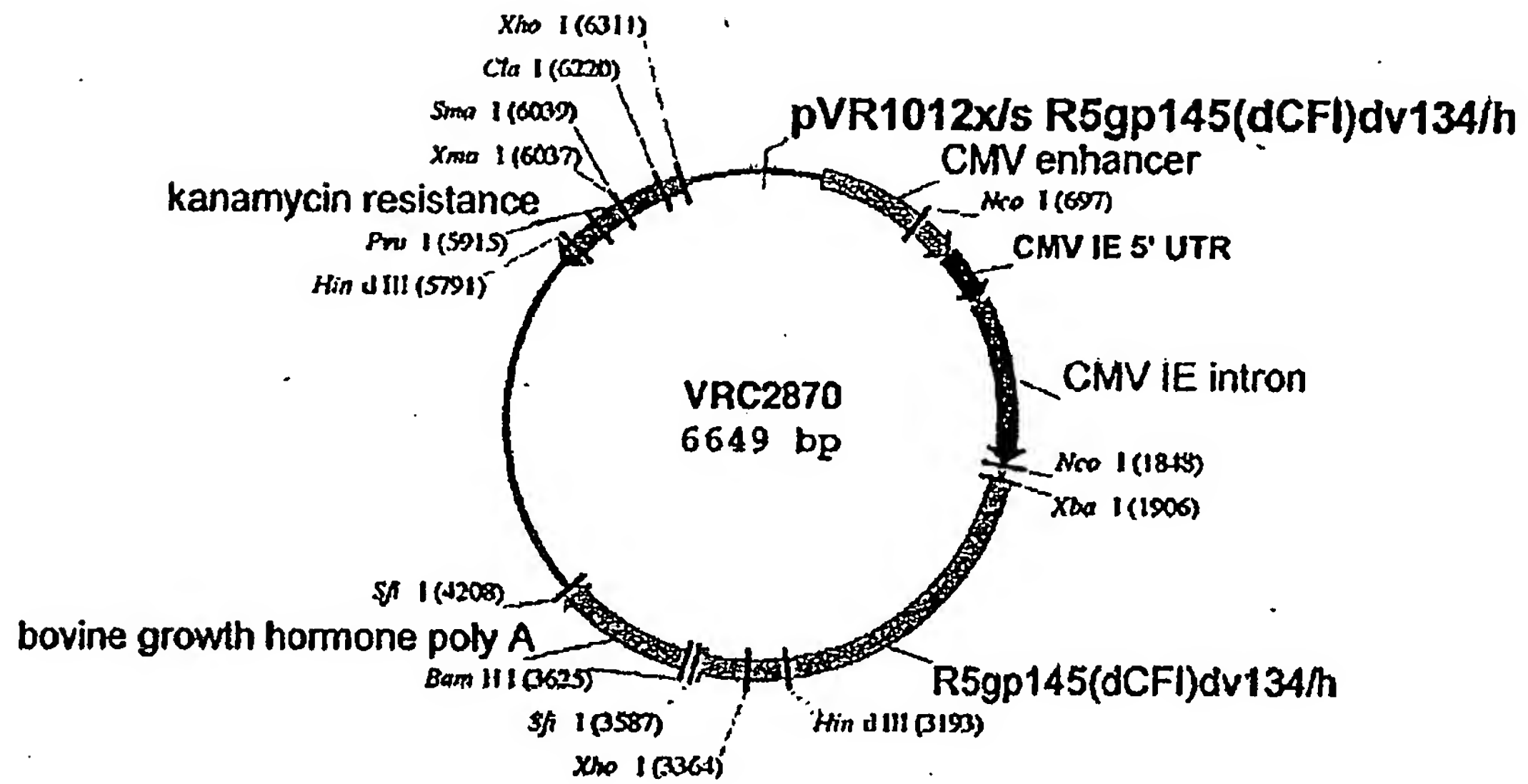


Figure 65



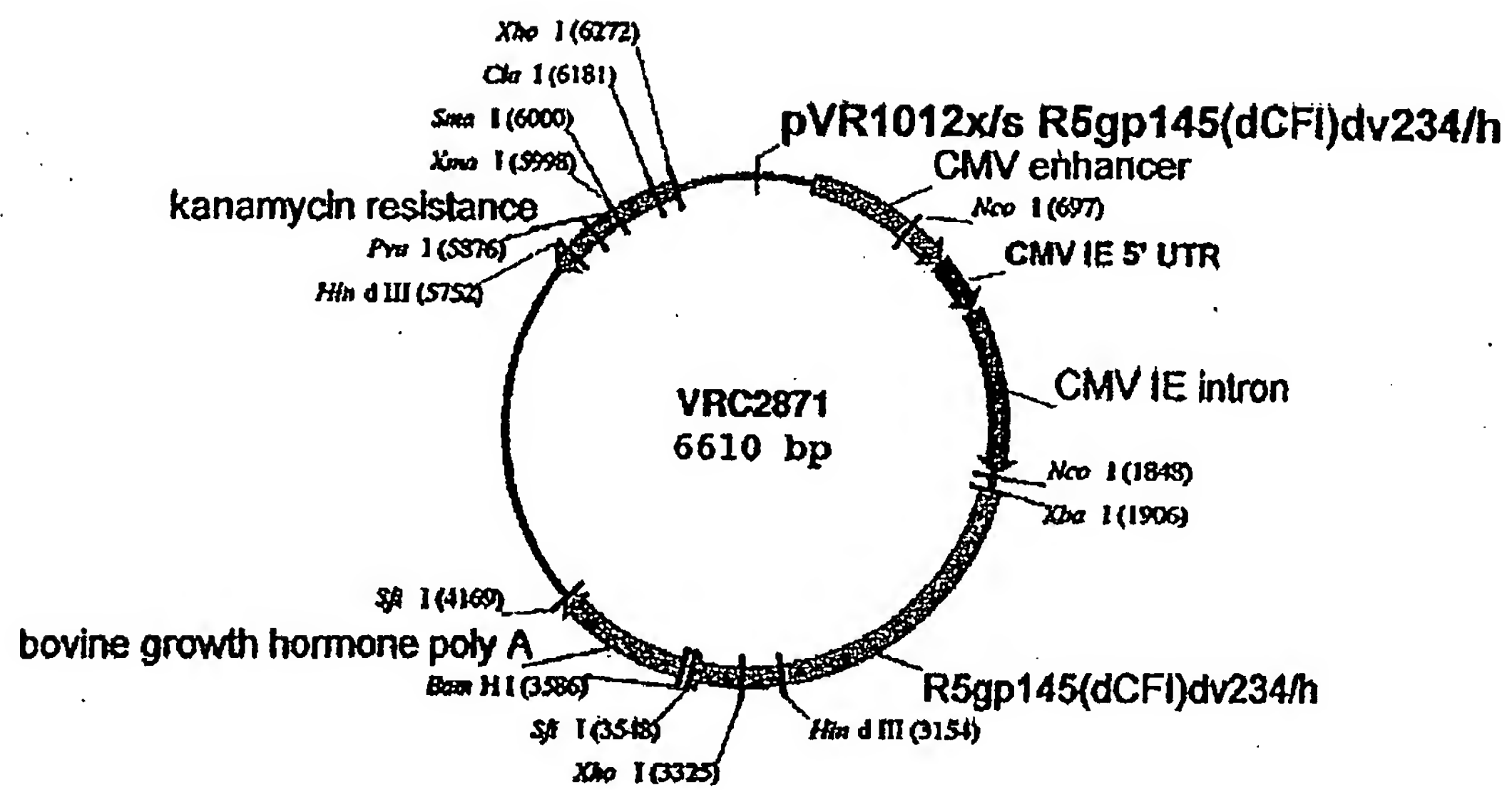


Figure 66

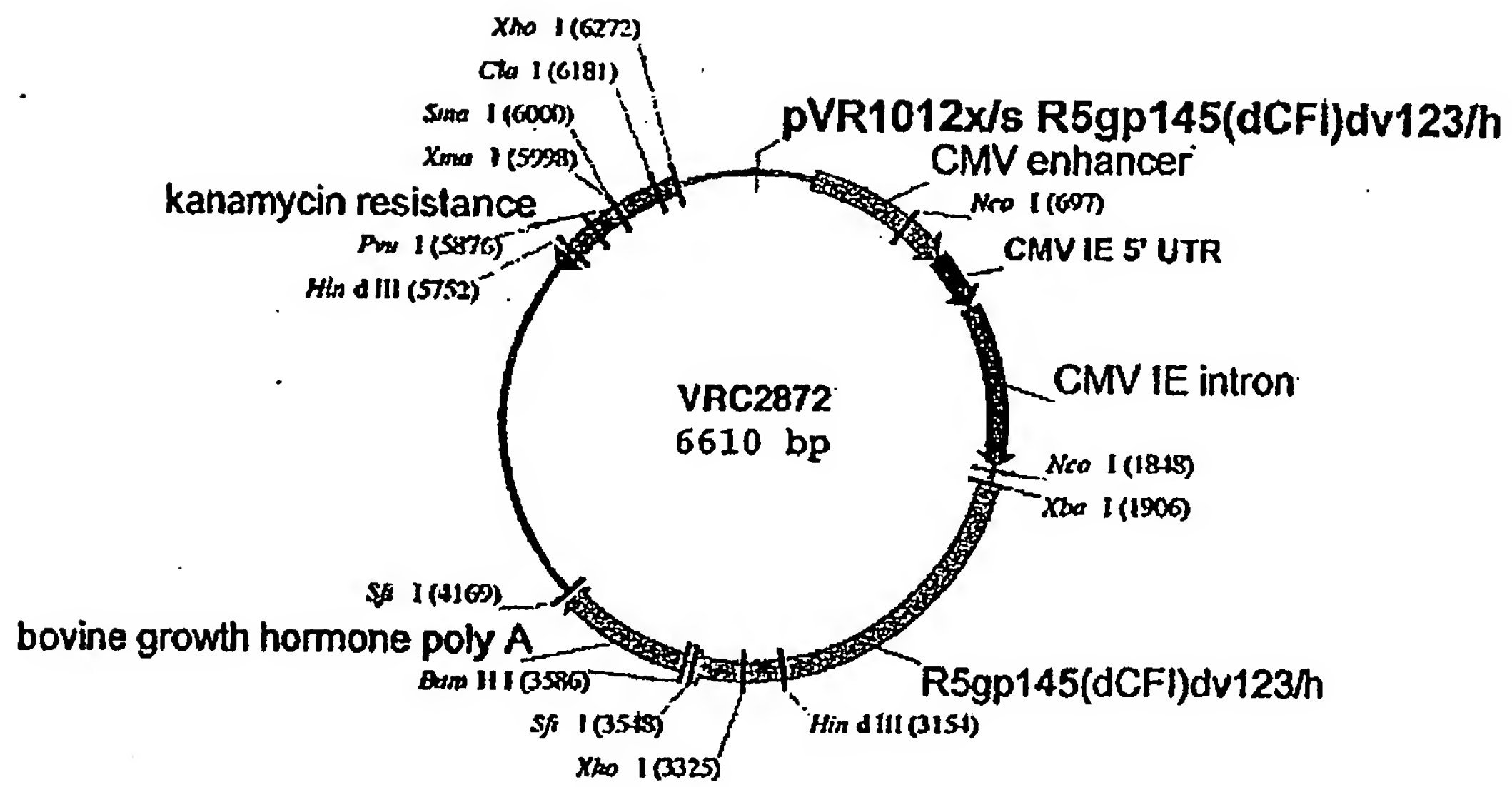


Figure 67

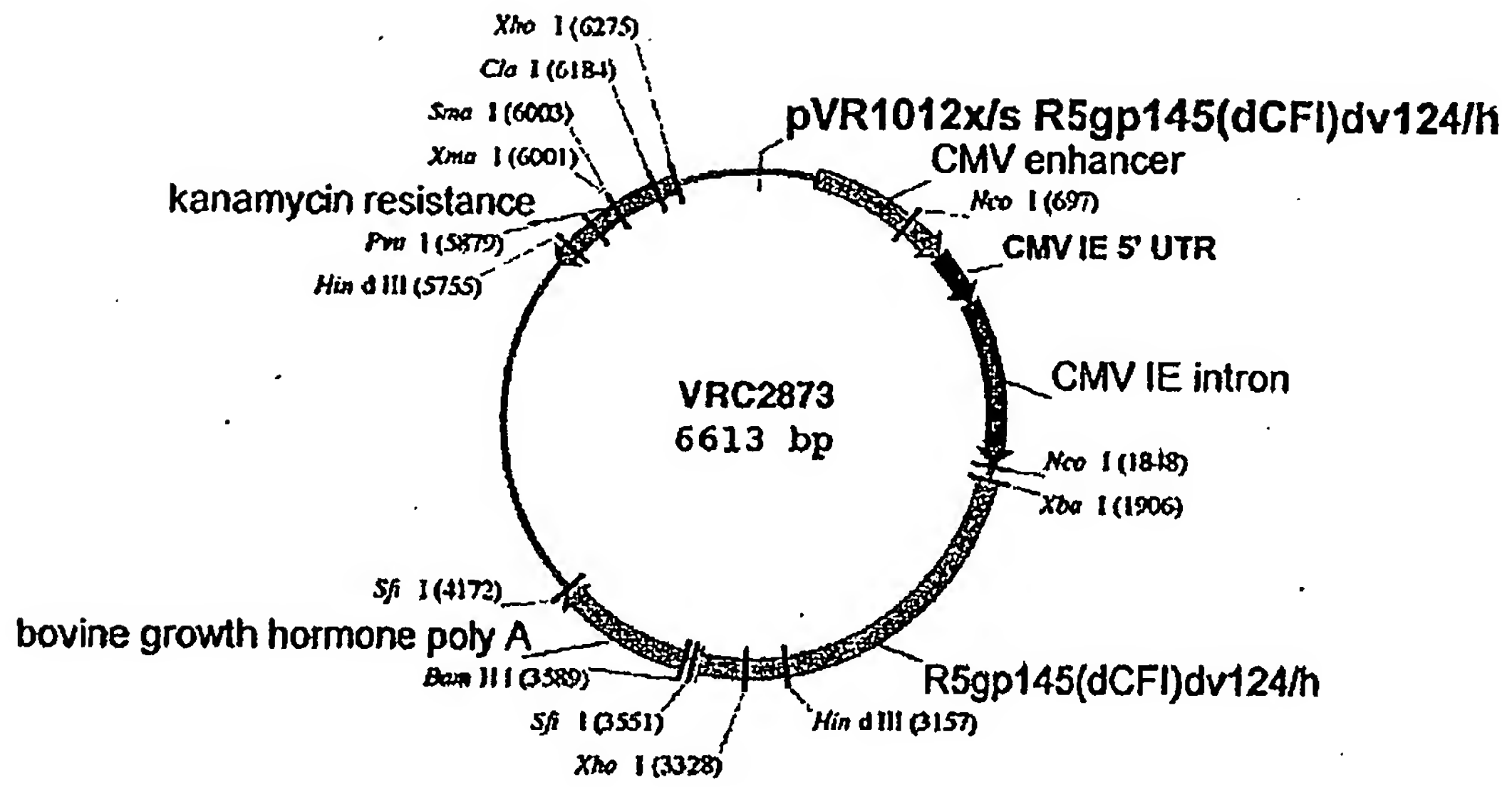


Figure 68

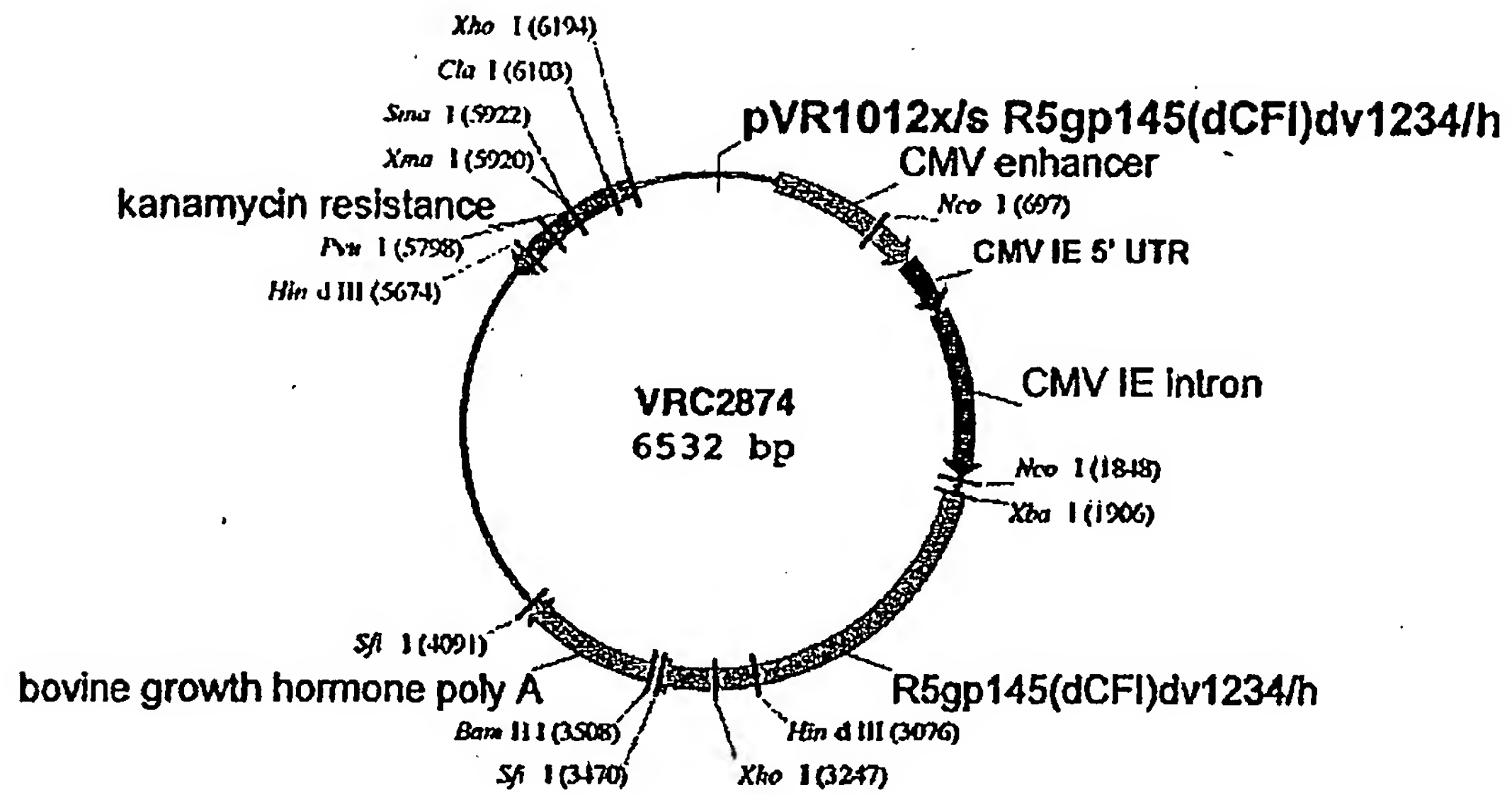


Figure 69

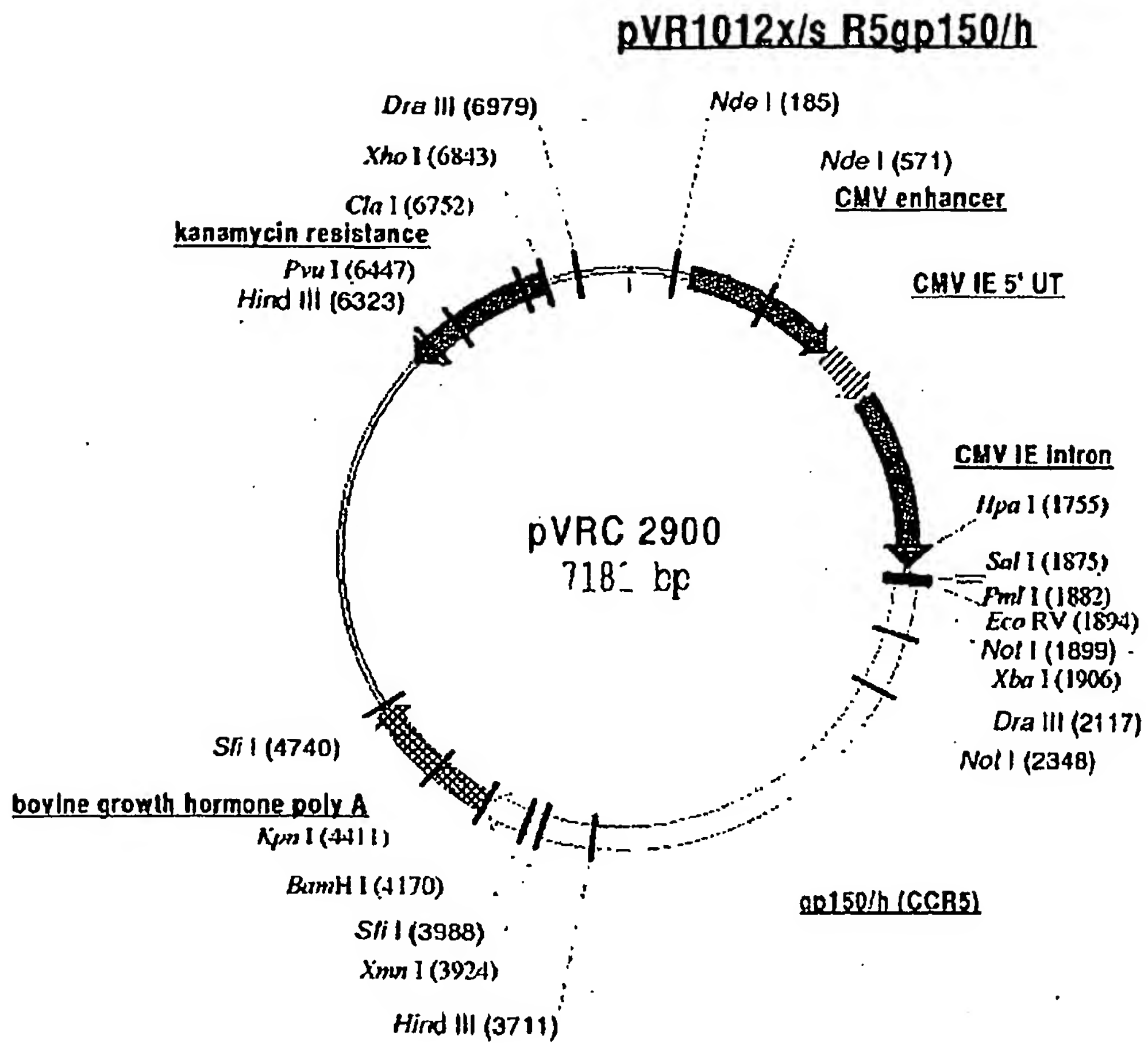


Figure 70

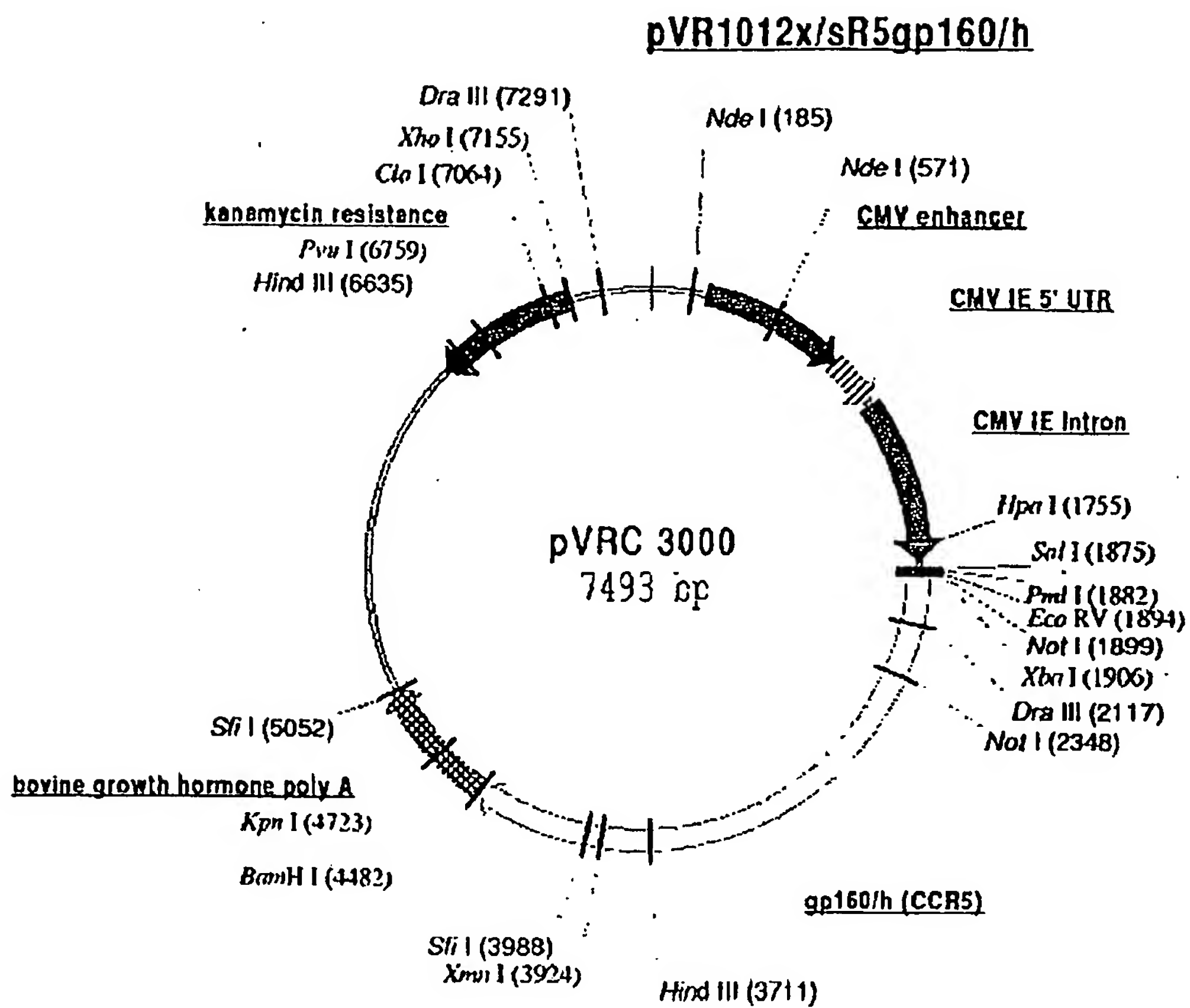


Figure 71



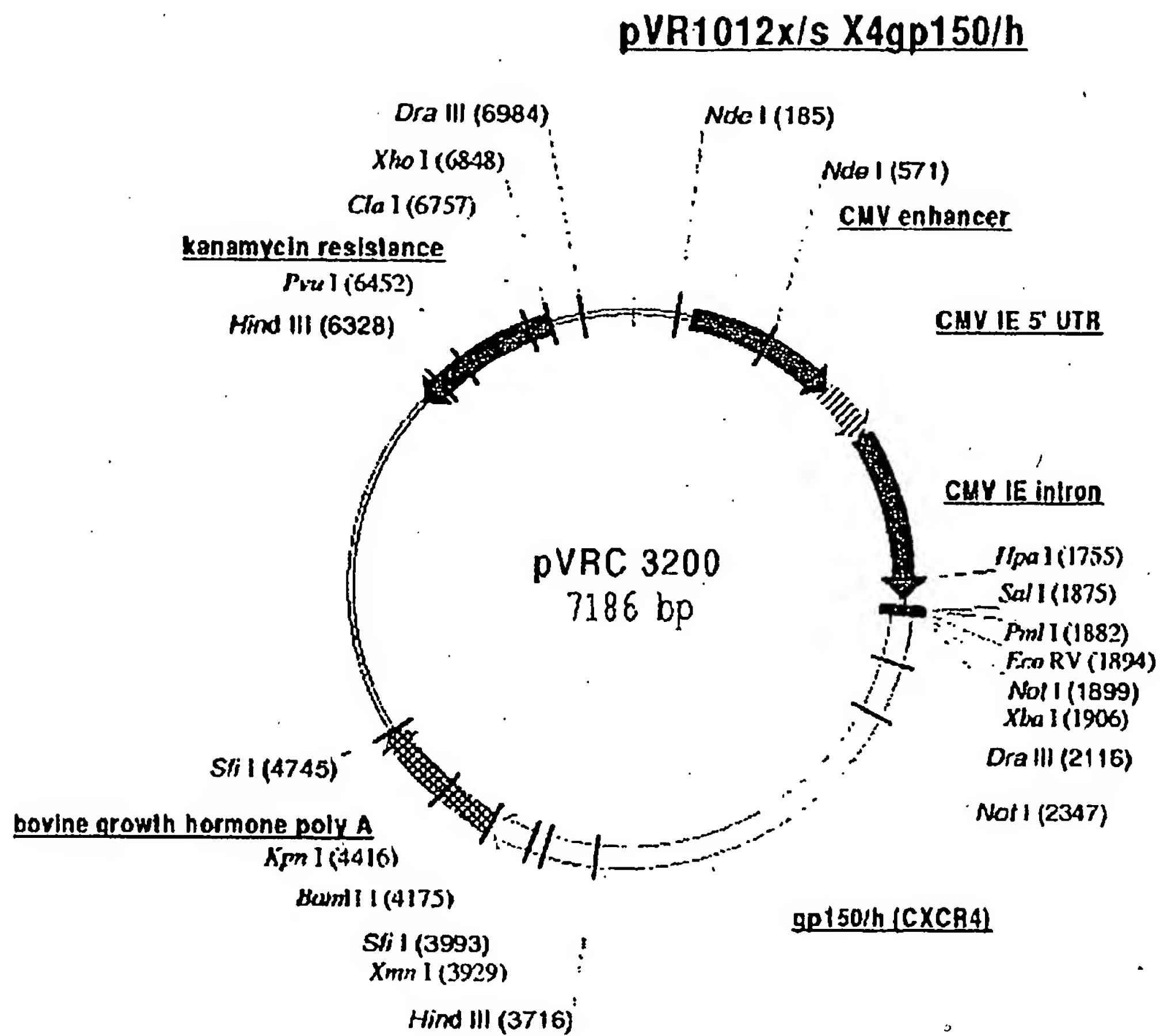


Figure 72

**pVR1012x/s X4gp150(del F/CL del HIS)/h**

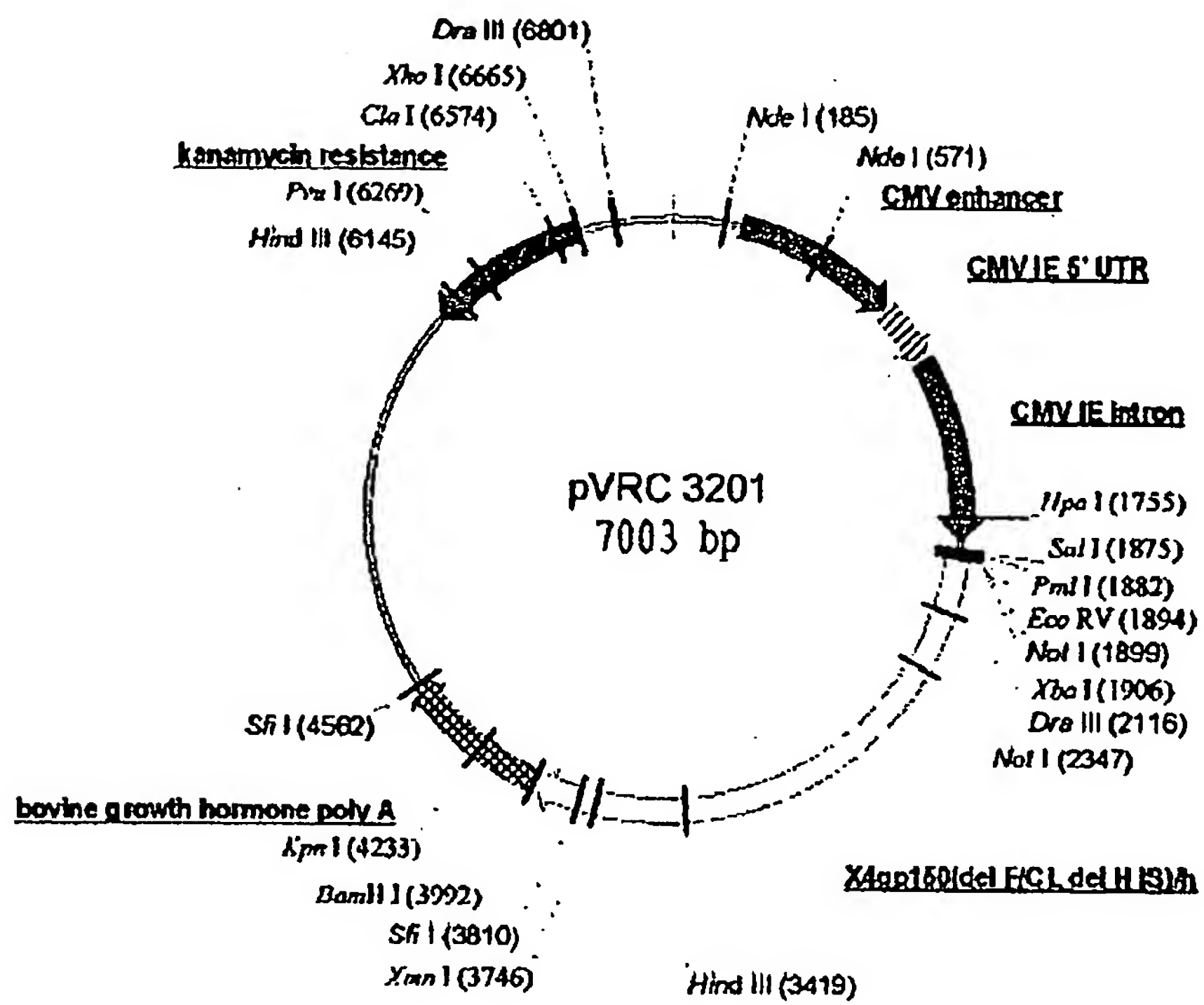


Figure 73

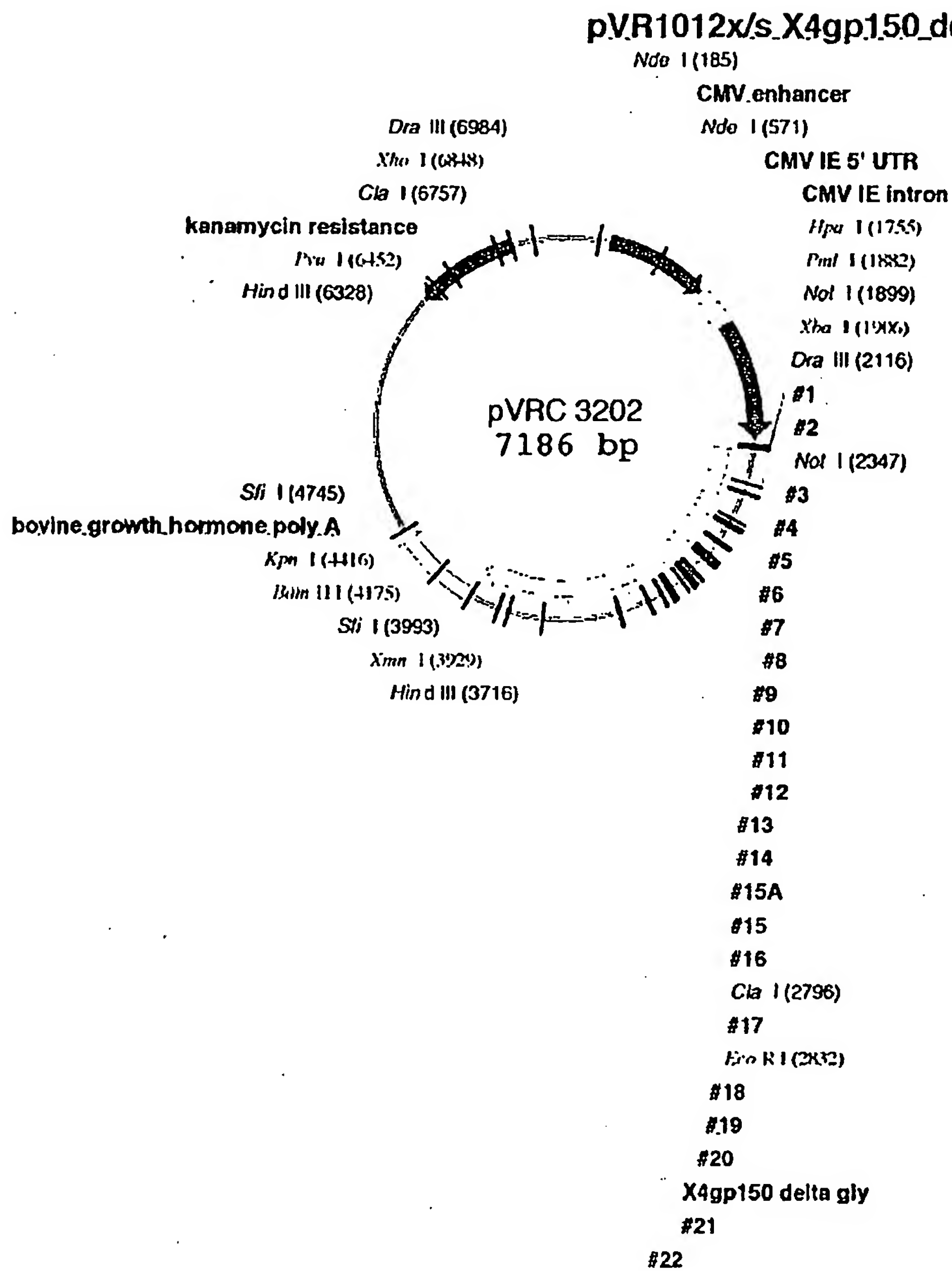


Figure 74

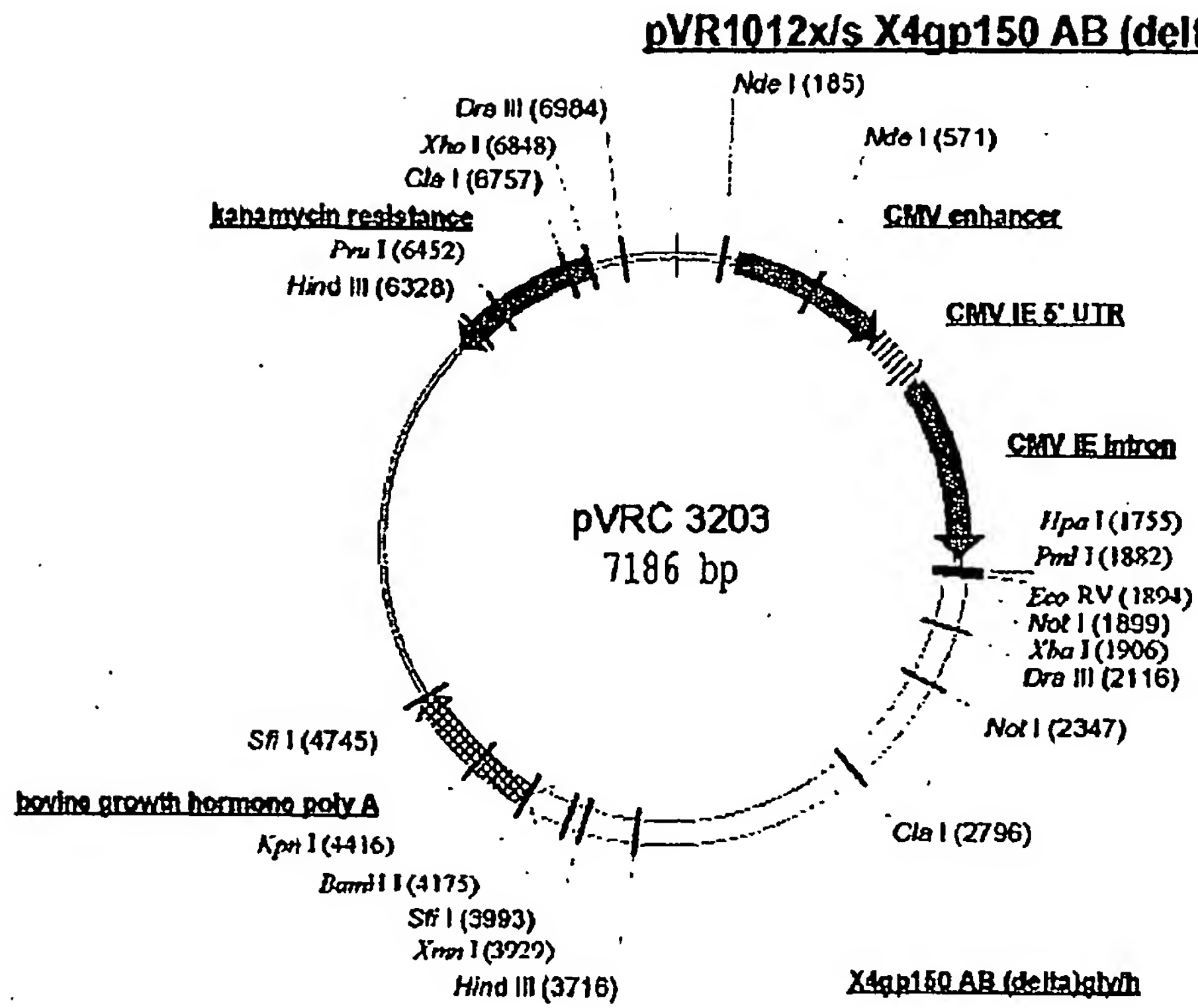


Figure 75

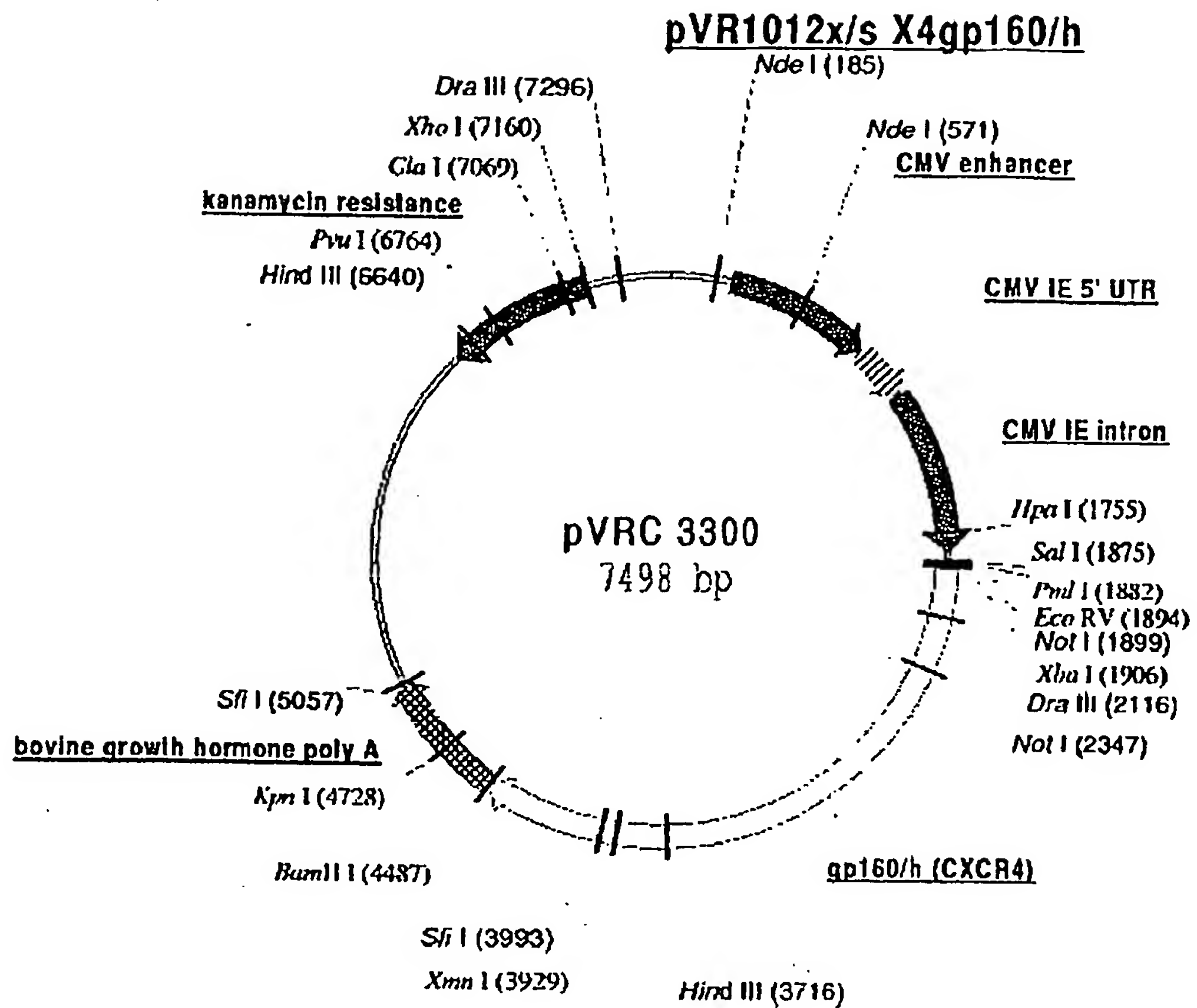


Figure 76

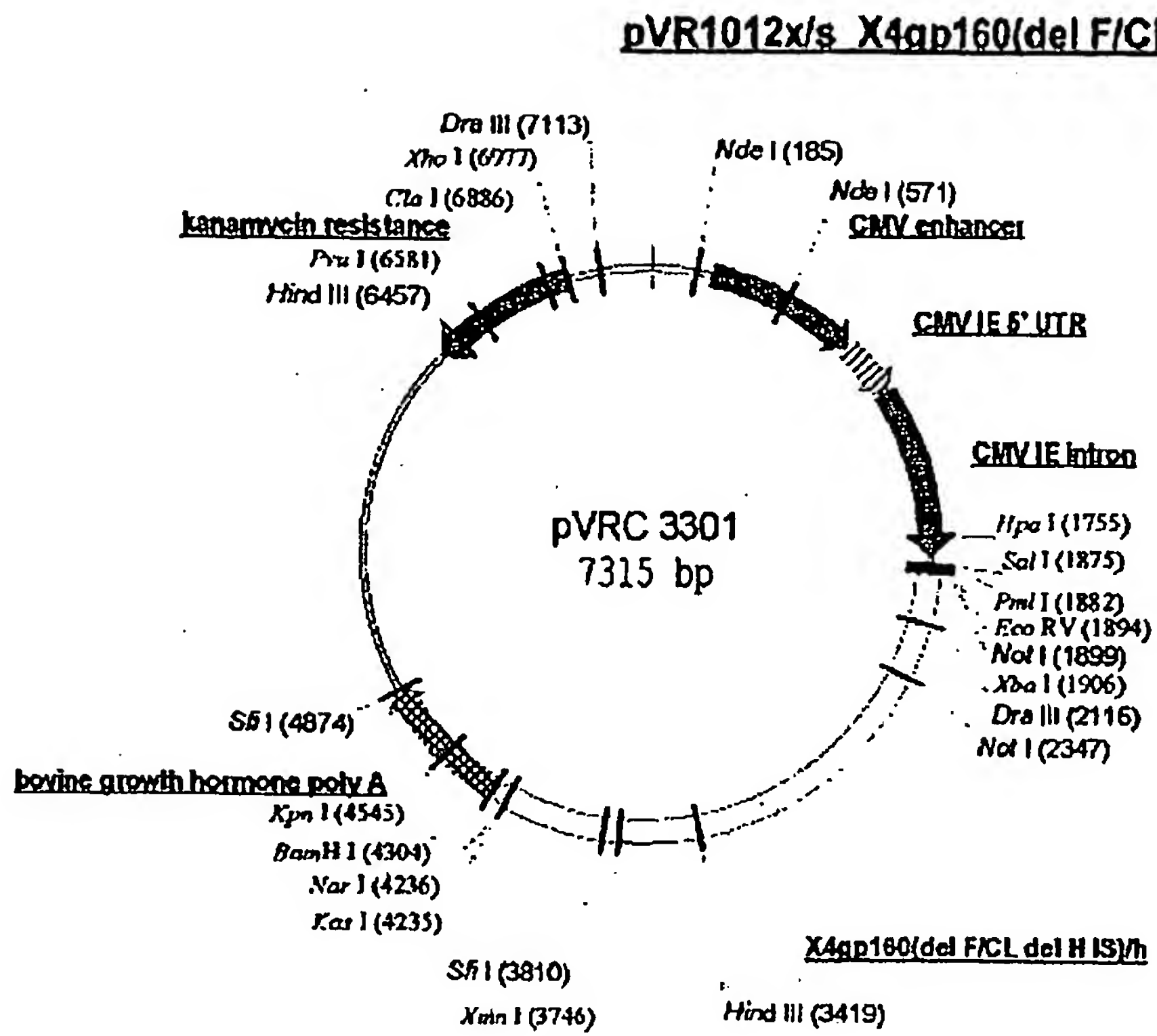


Figure 77



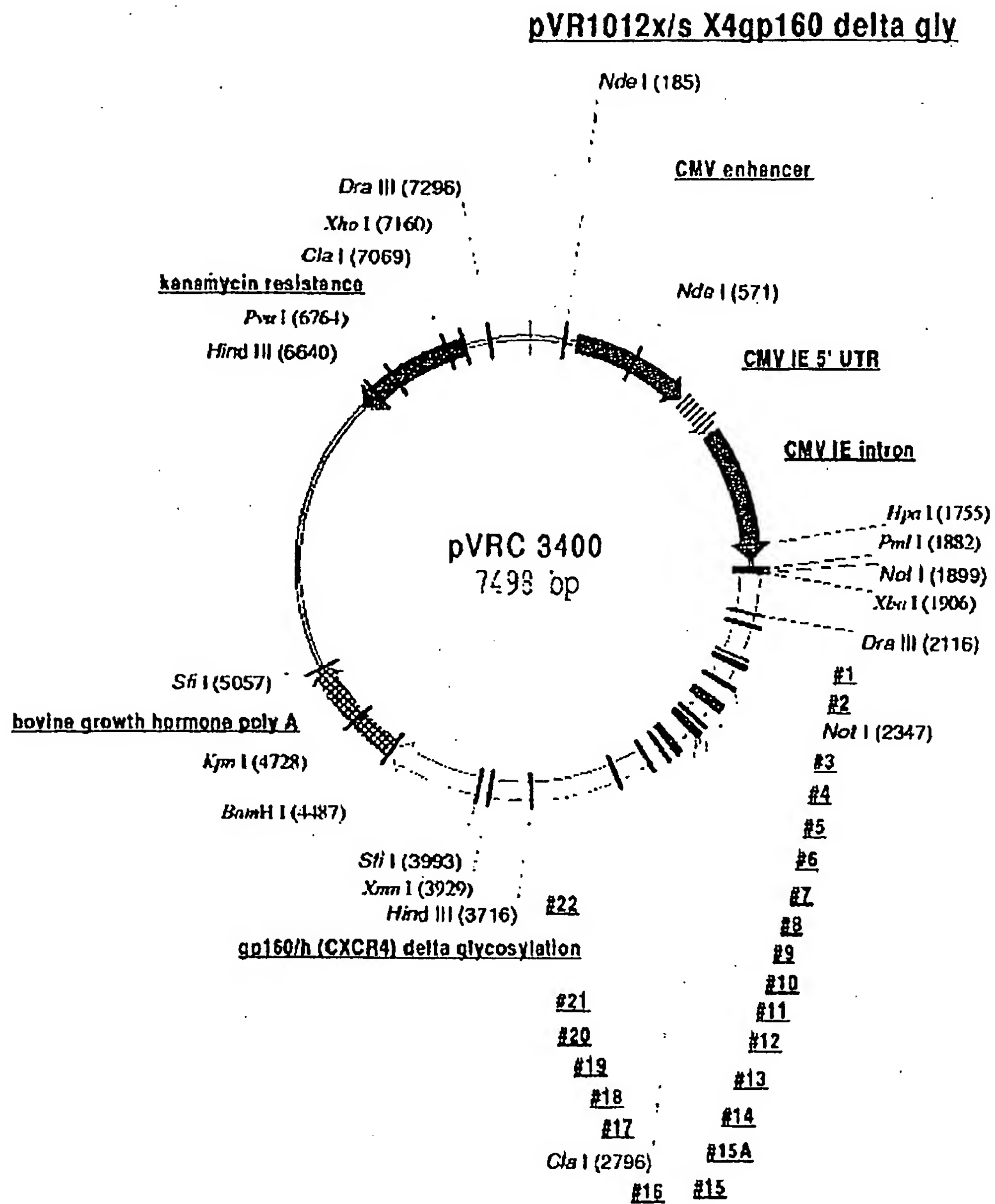


Figure 78

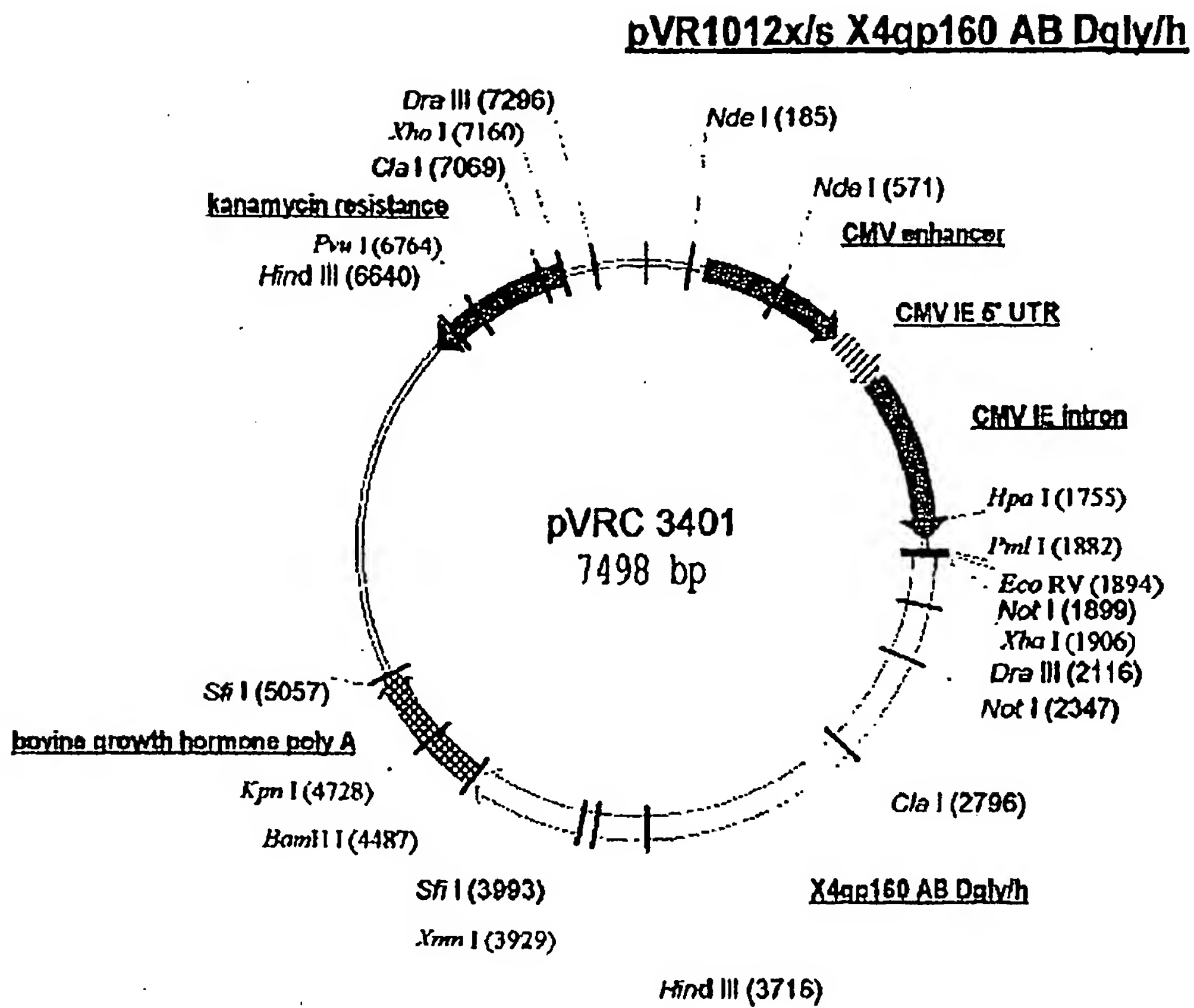


Figure 79

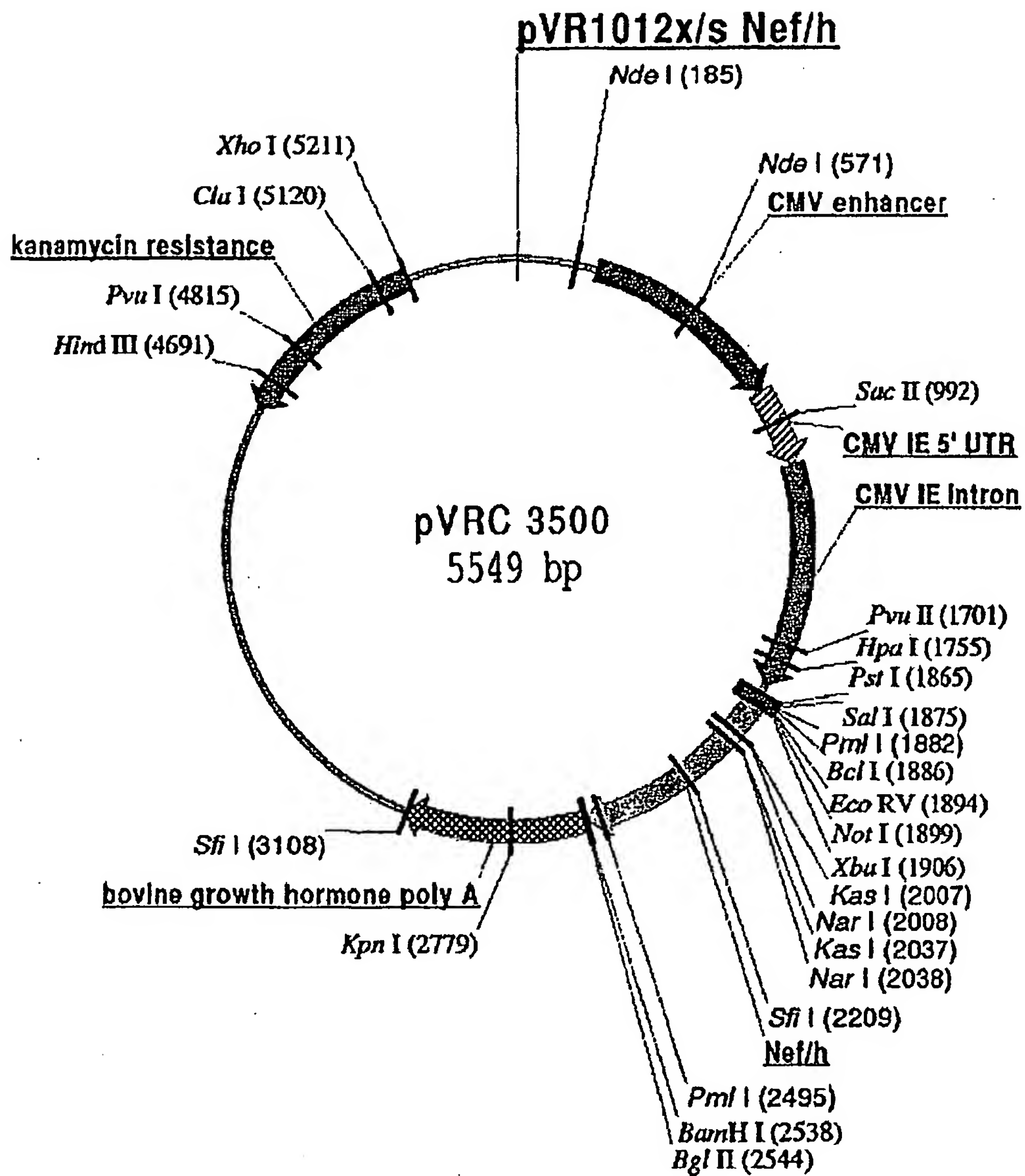


Figure 80

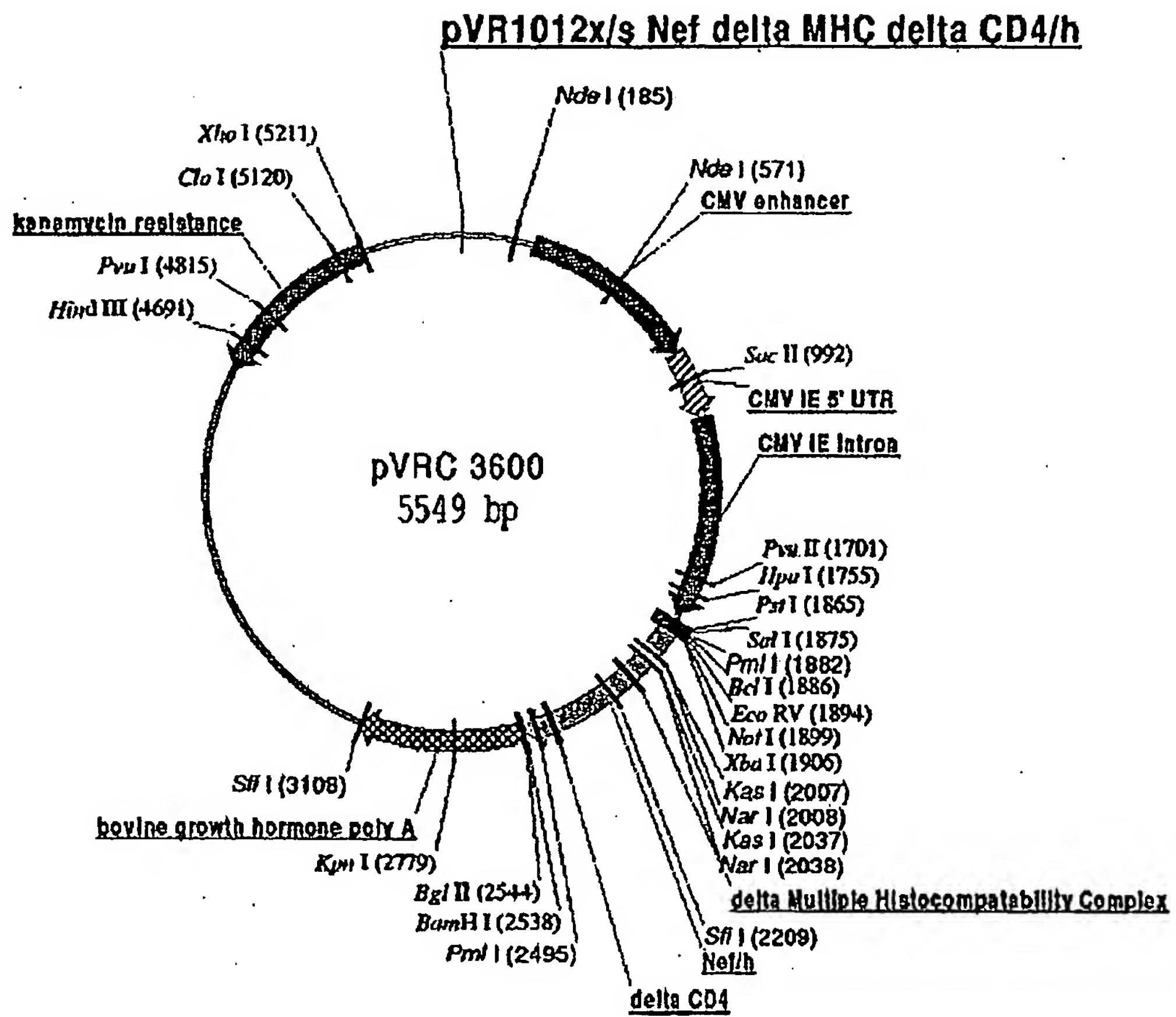


Figure 81

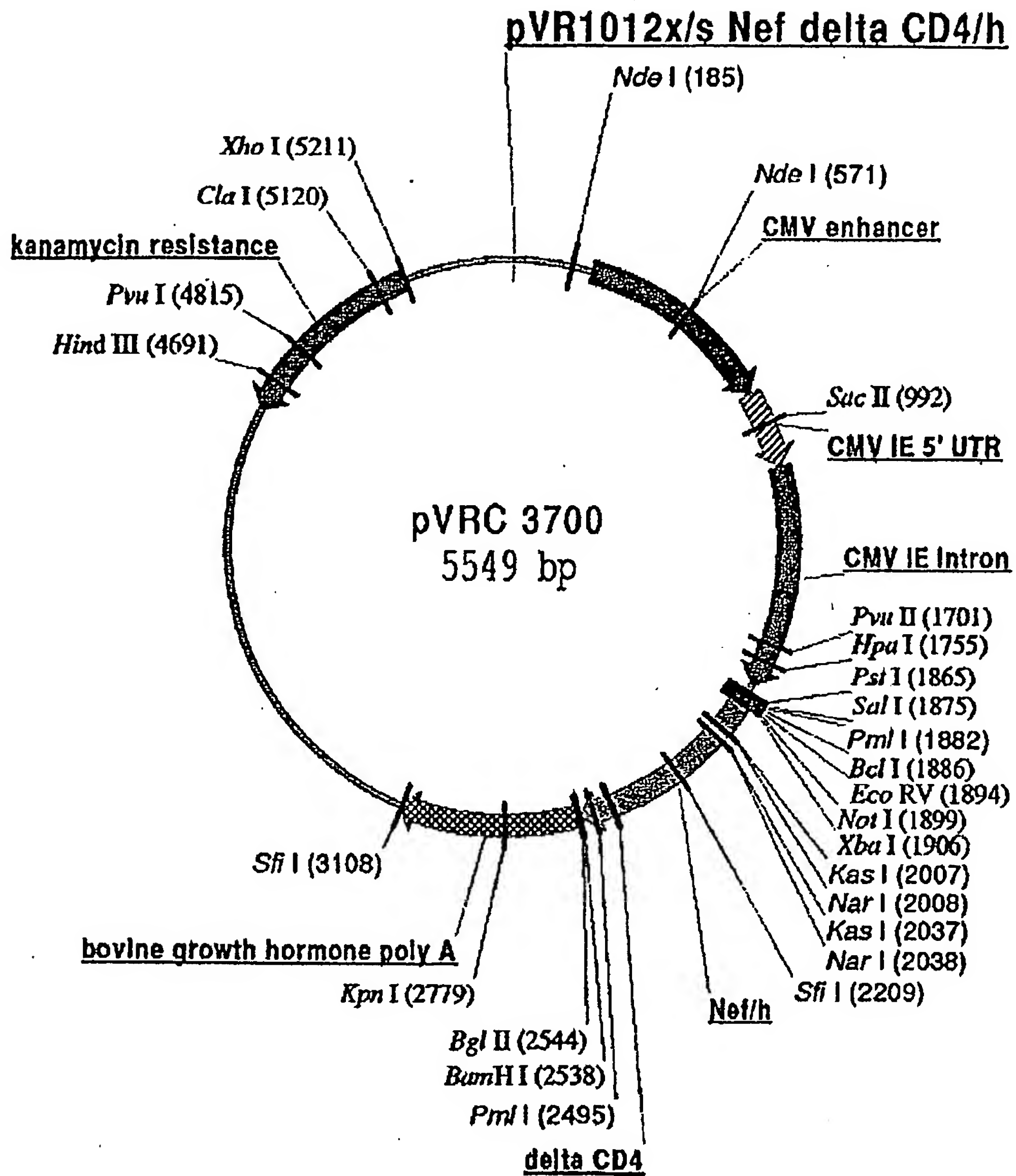


Figure 82

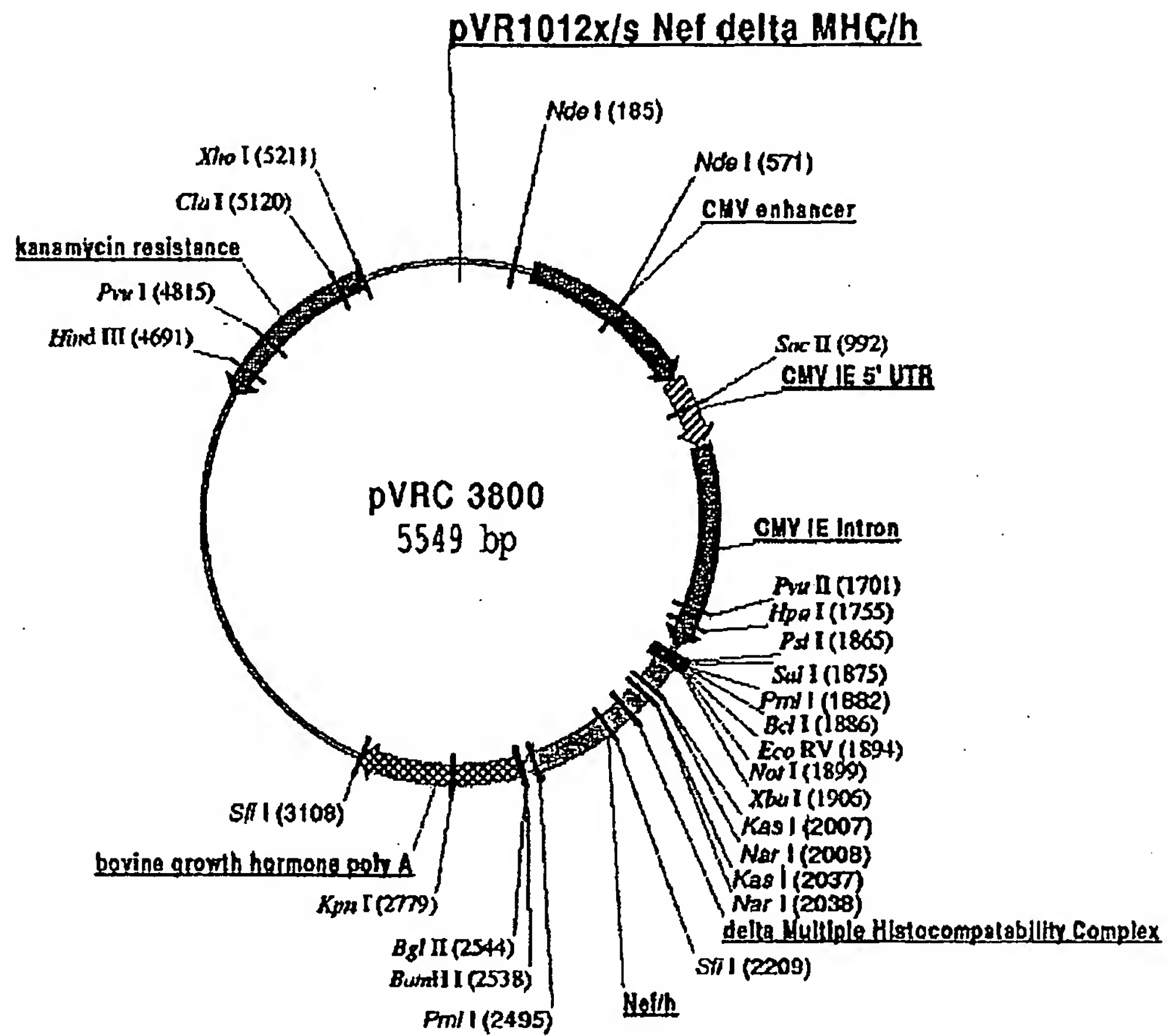


Figure 83



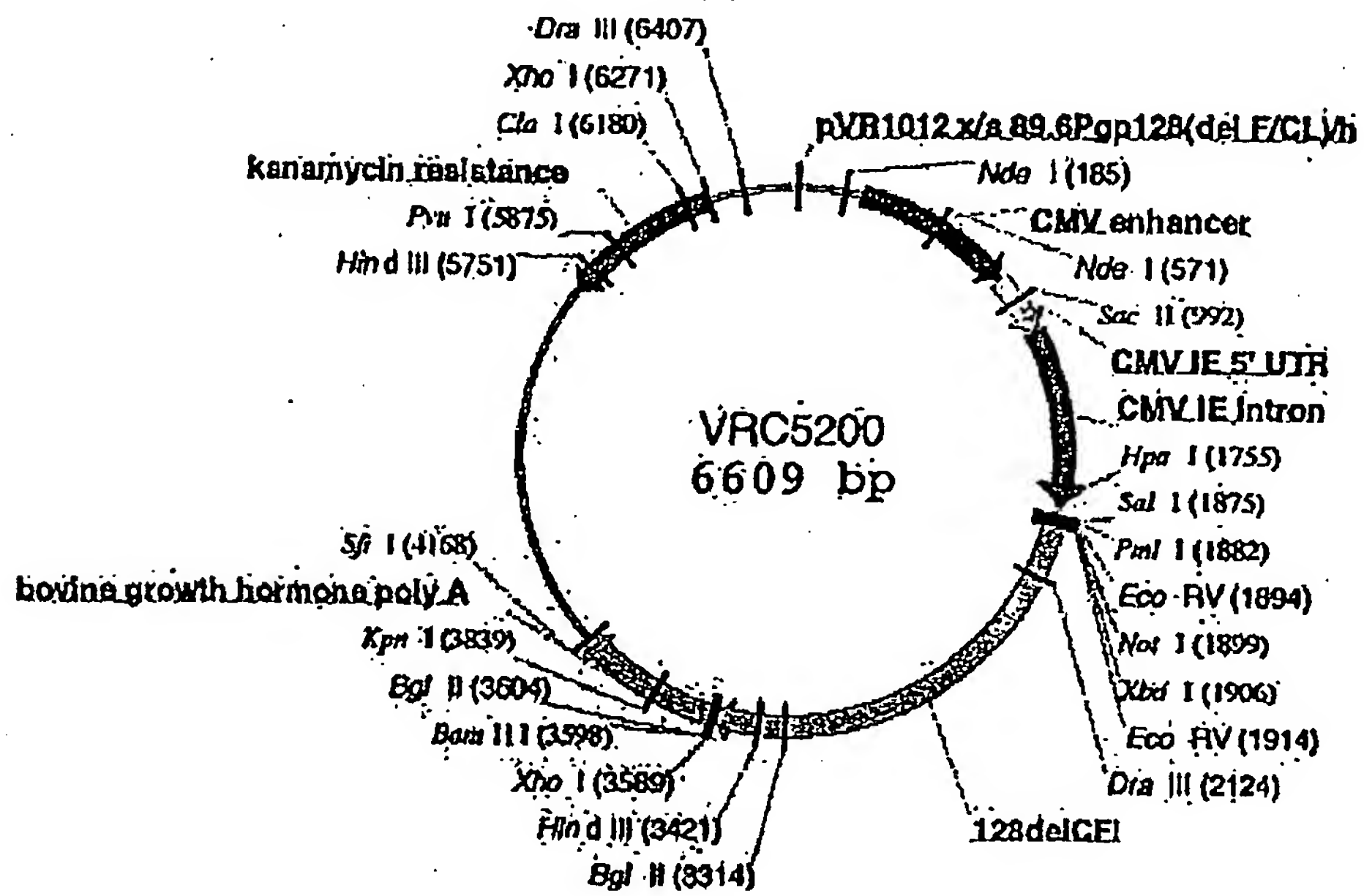


Figure 84

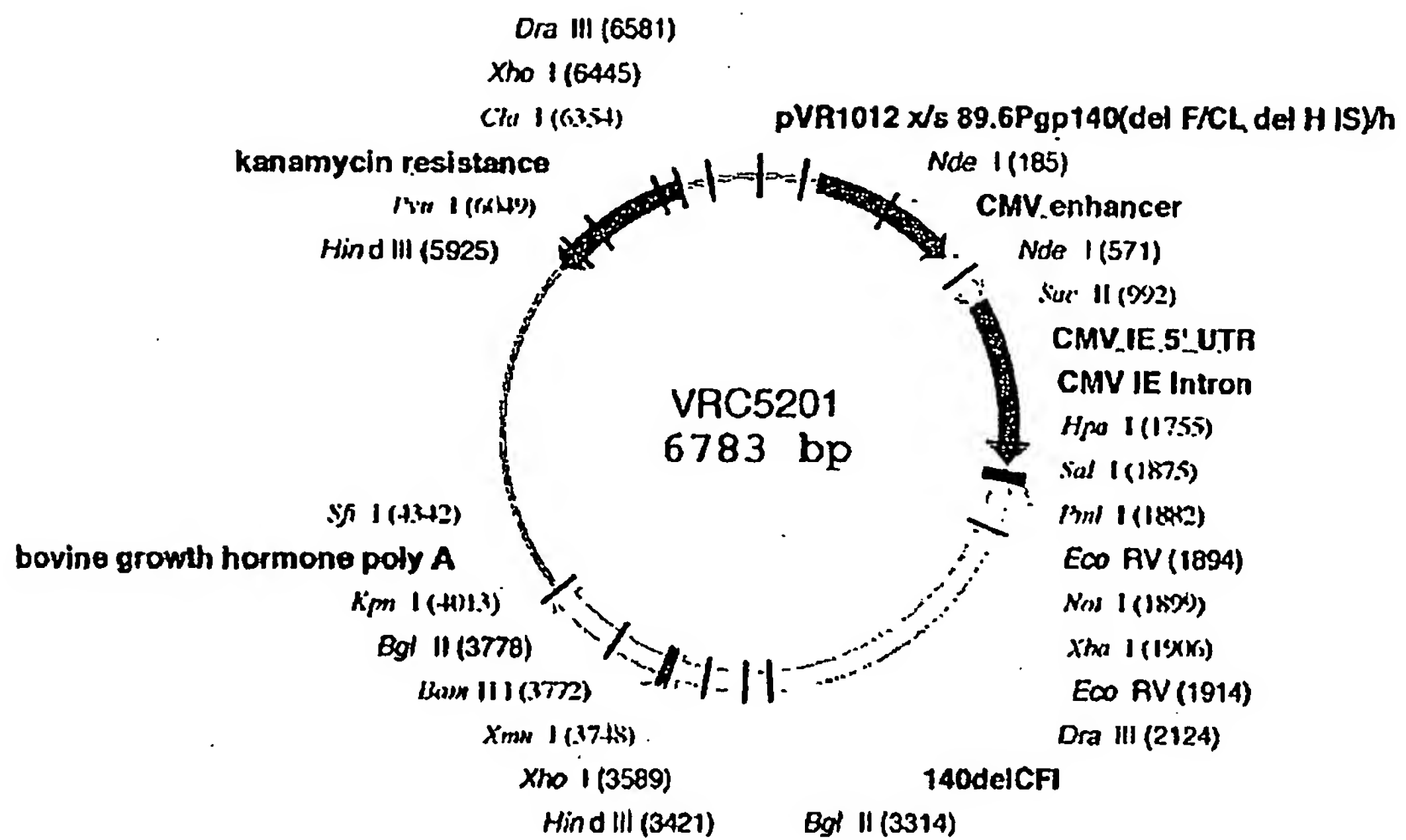


Figure 85

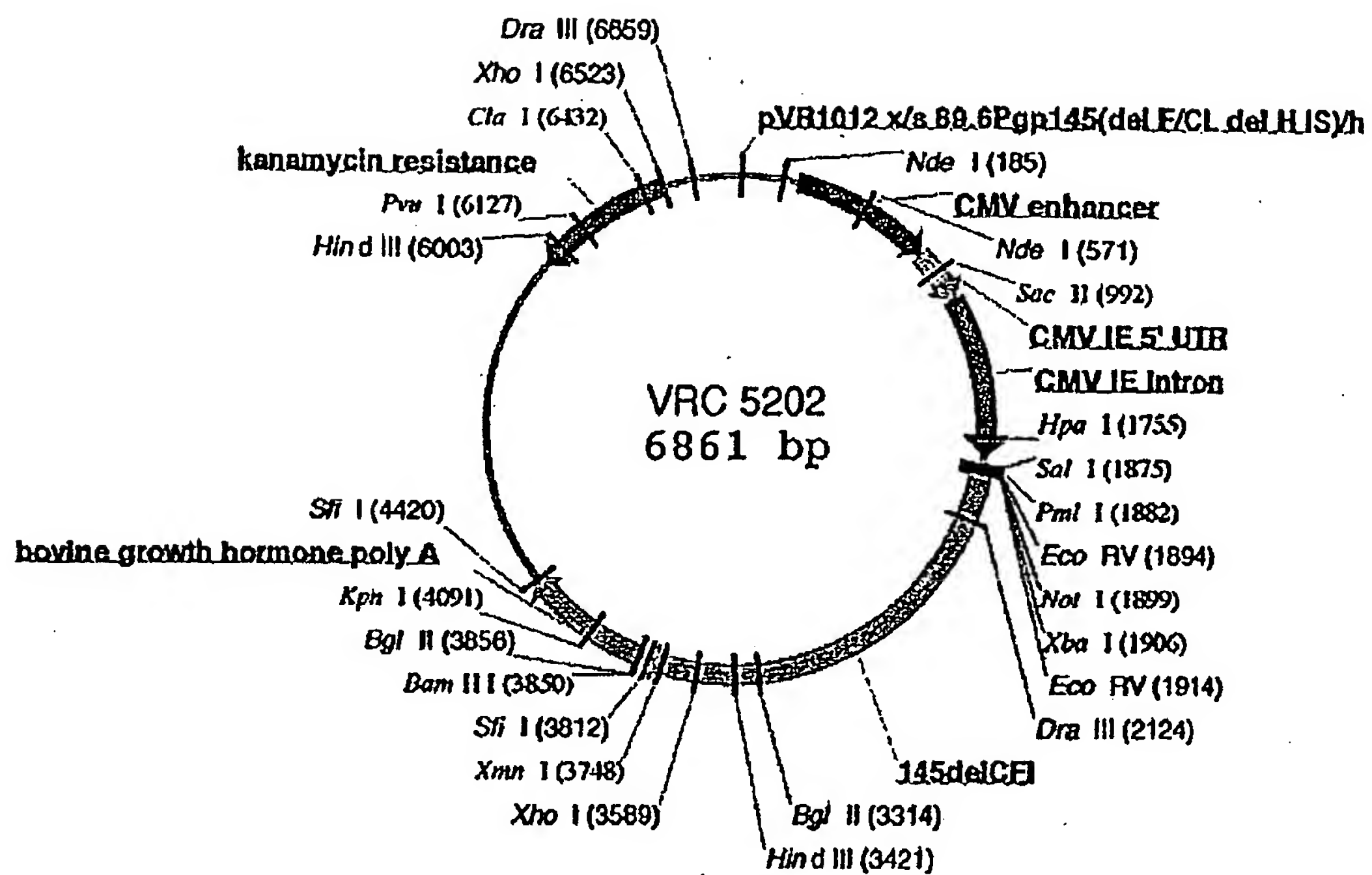


Figure 86

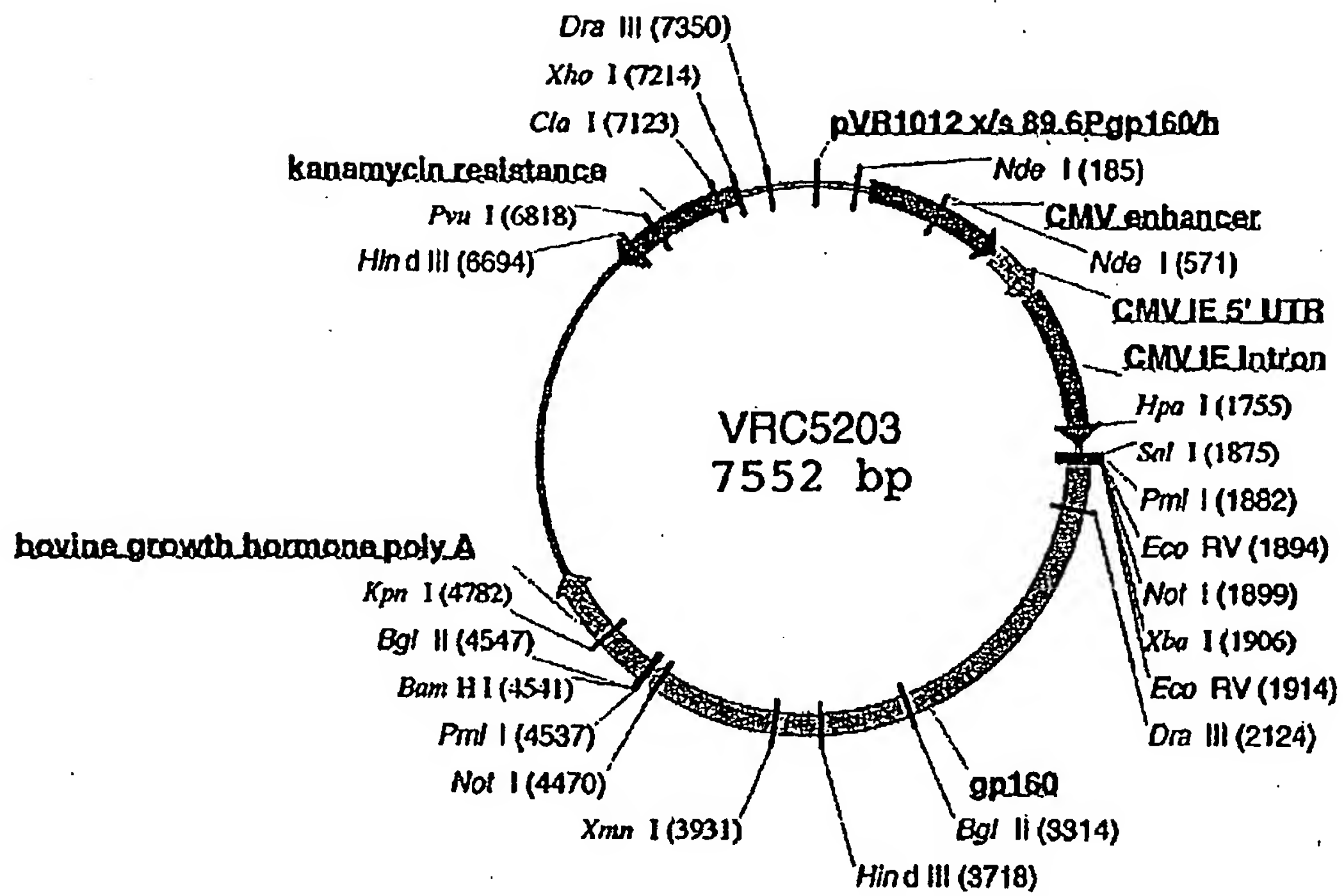


Figure 87

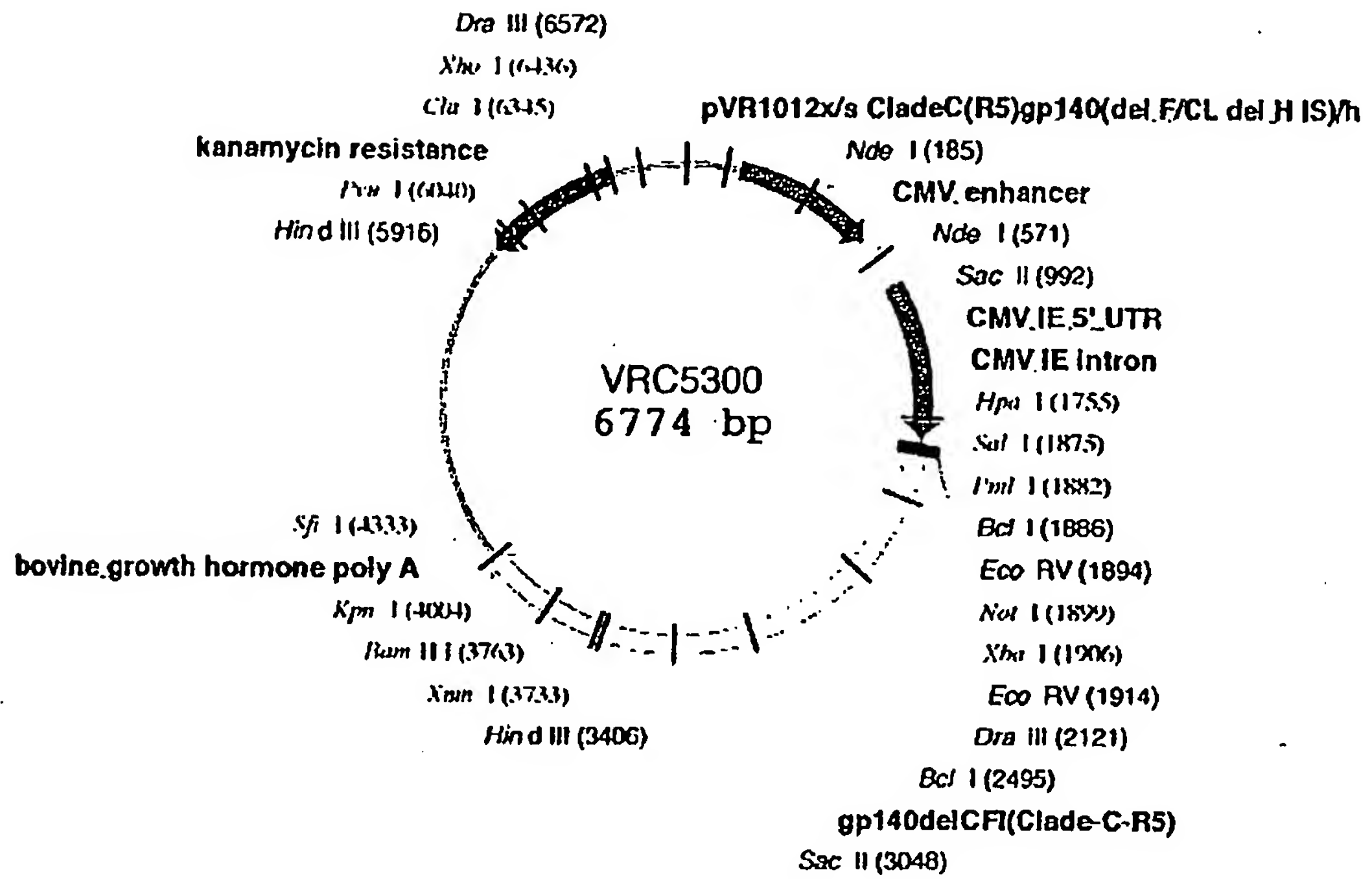


Figure 88

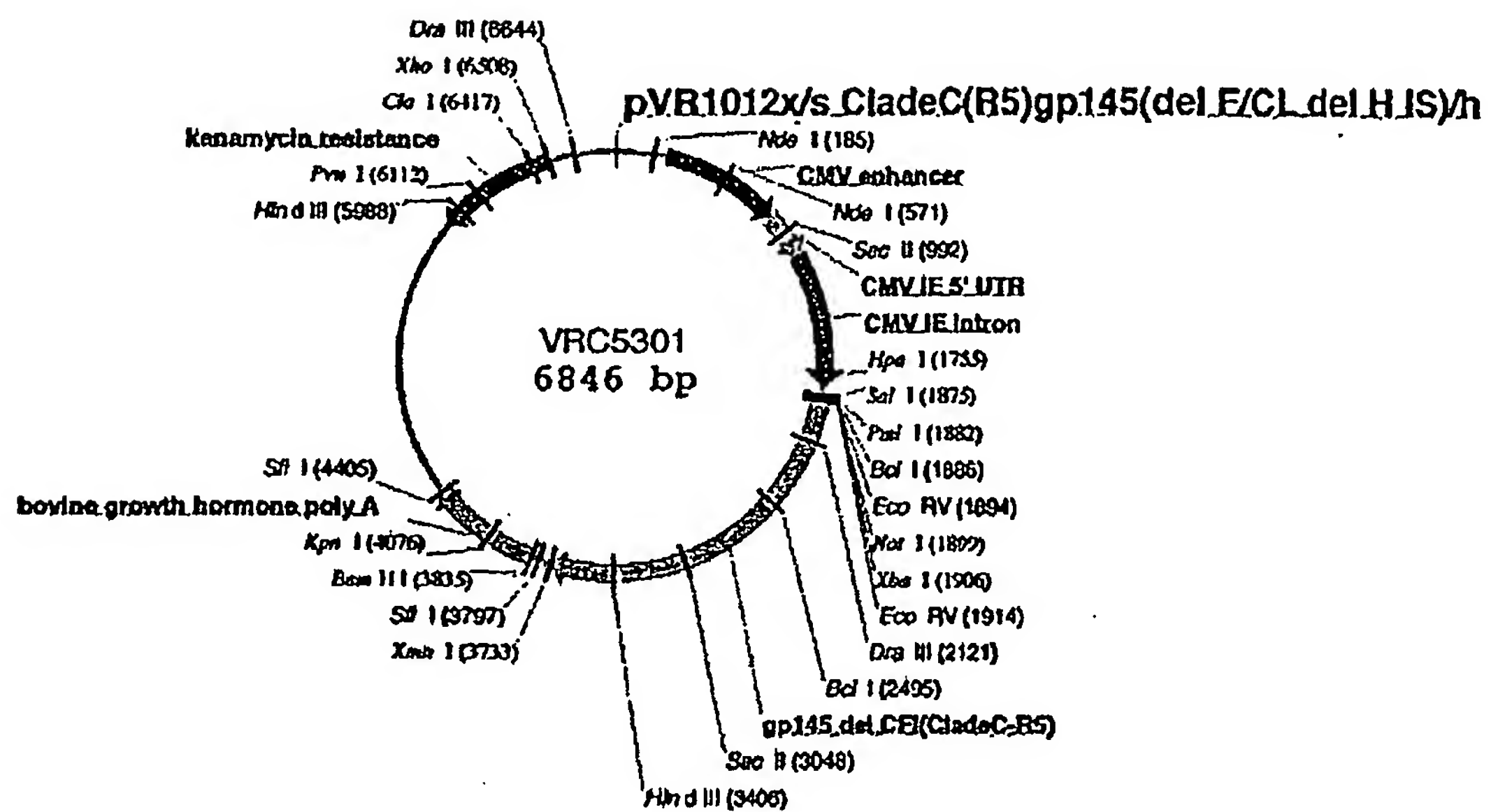


Figure 89

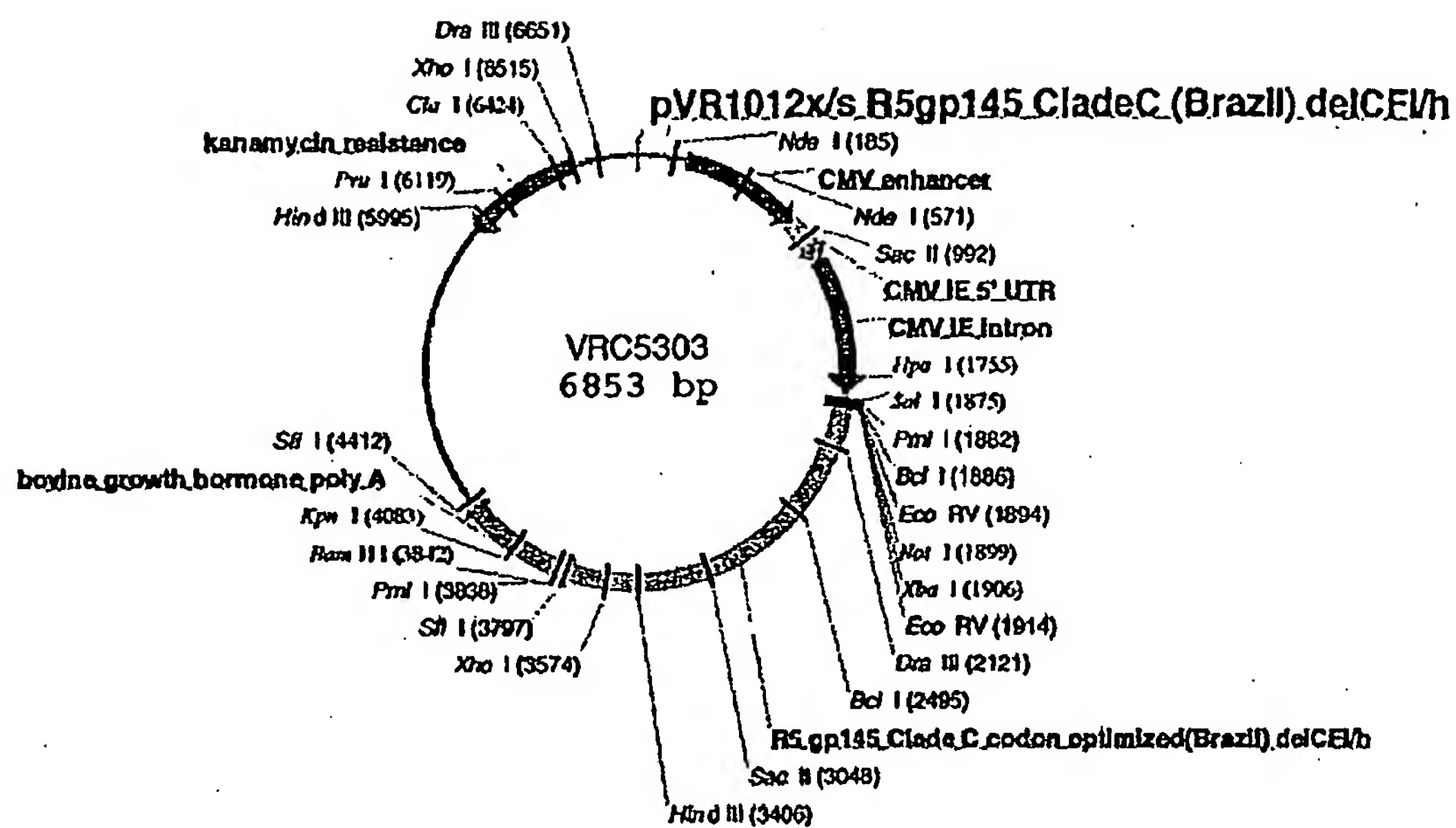


Figure 90



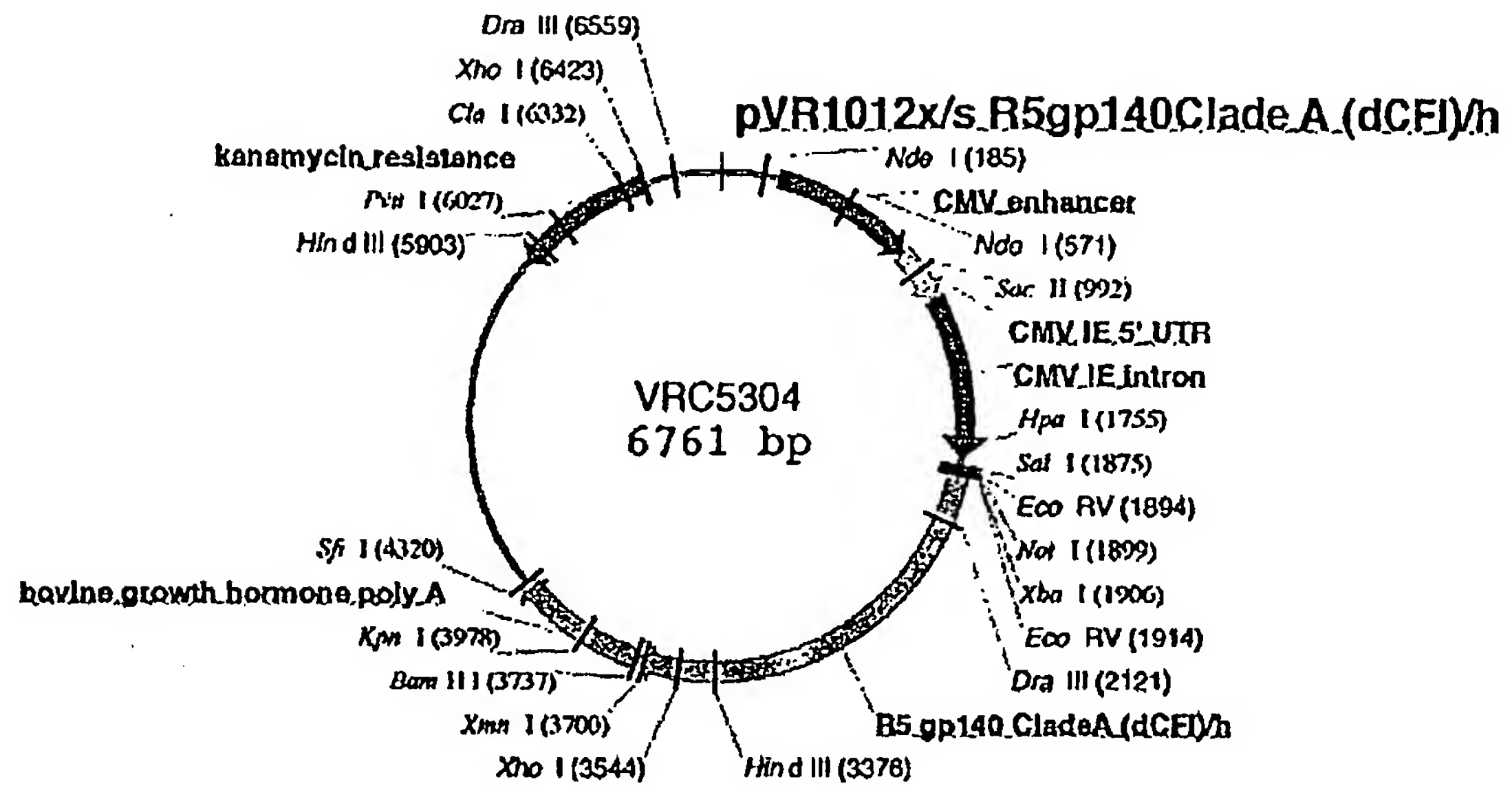


Figure 91

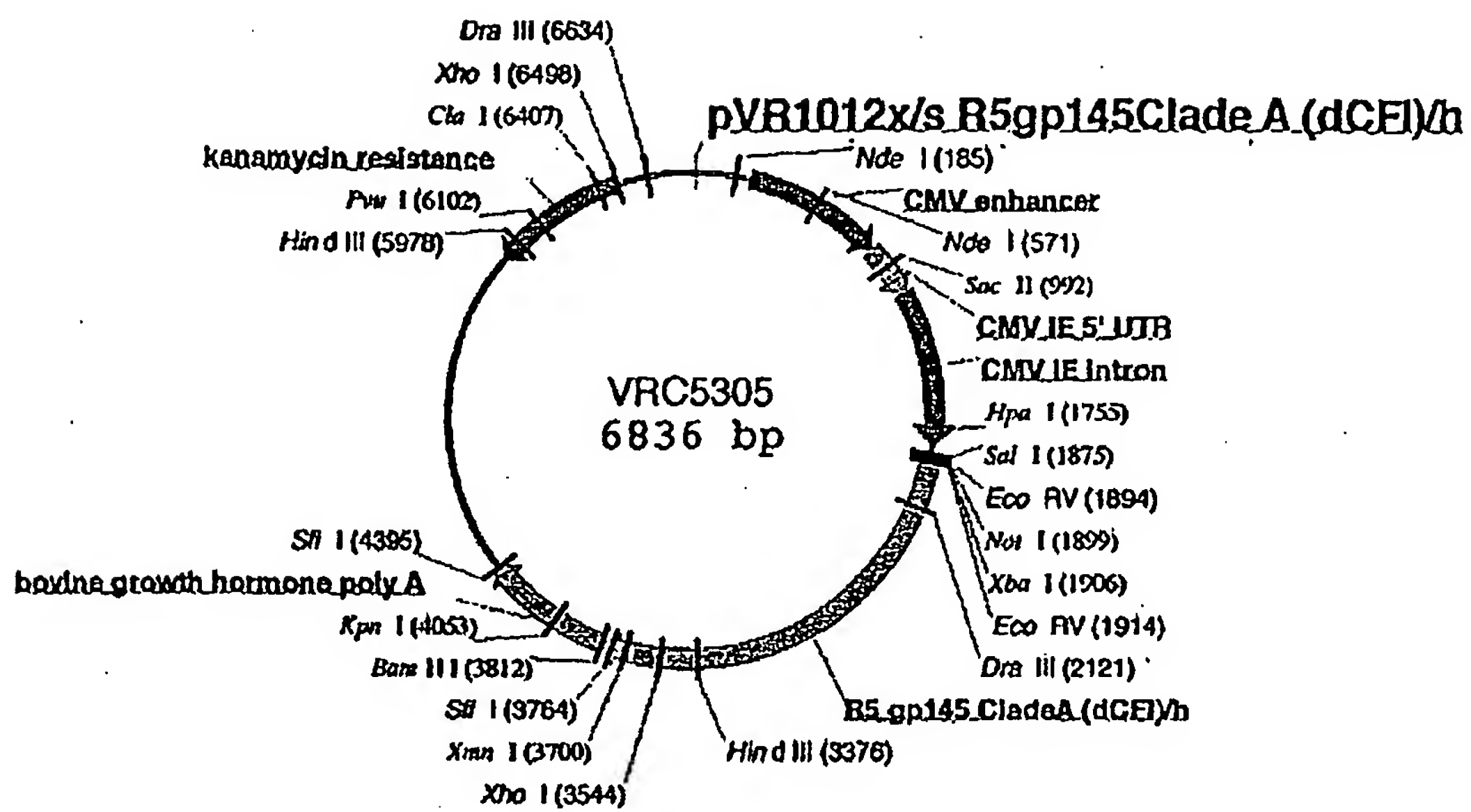


Figure 92

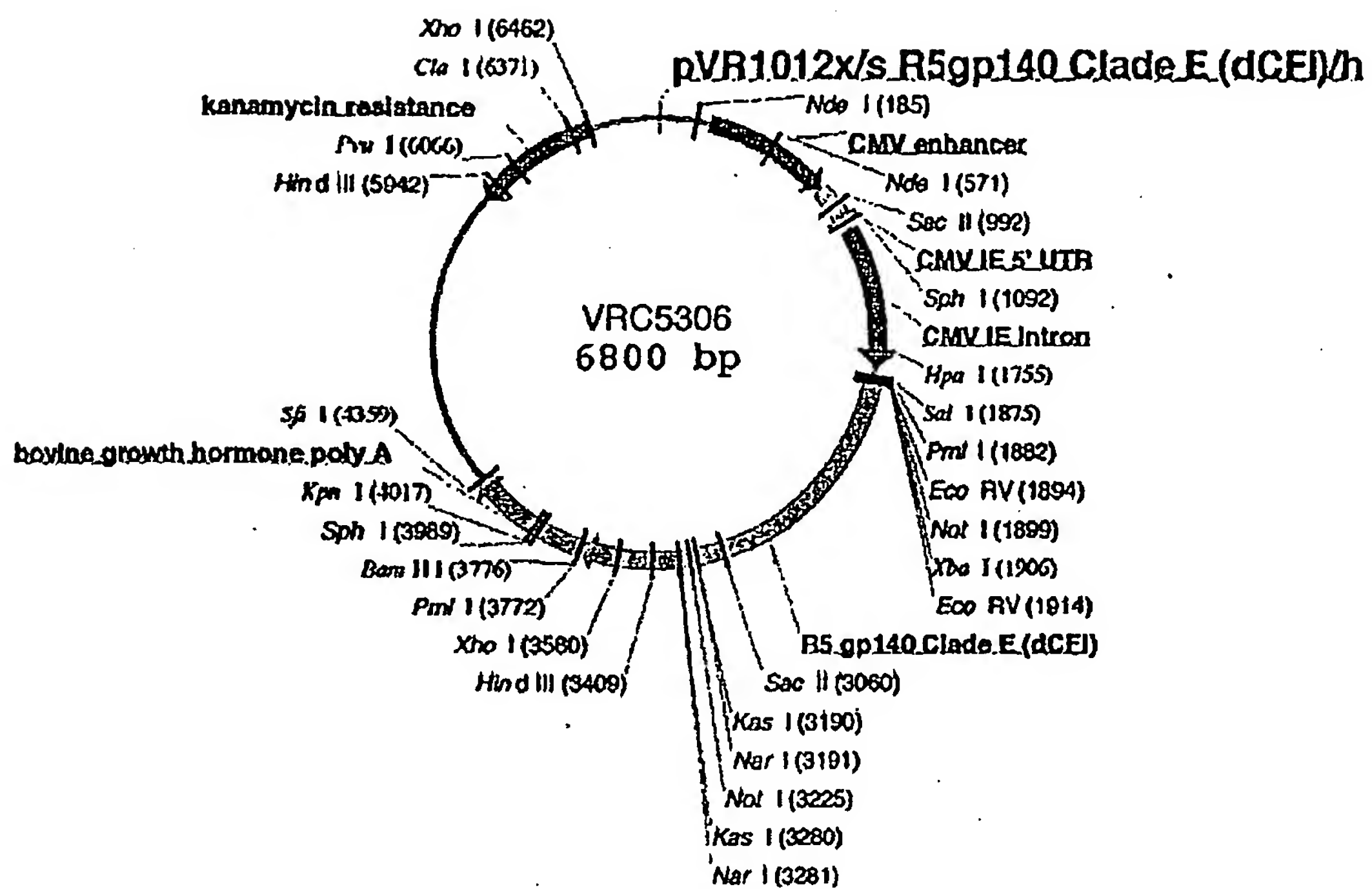


Figure 93

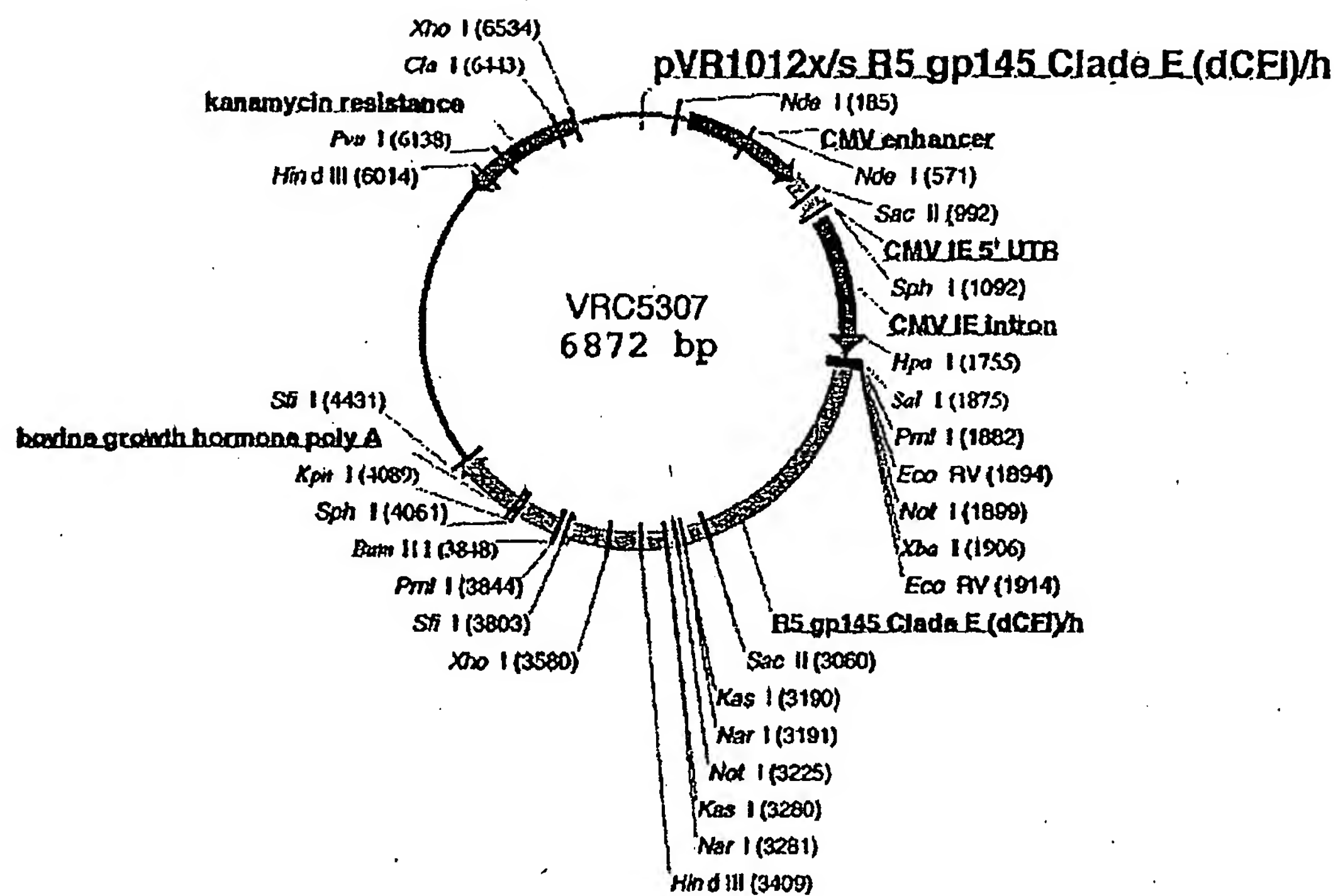


Figure 94

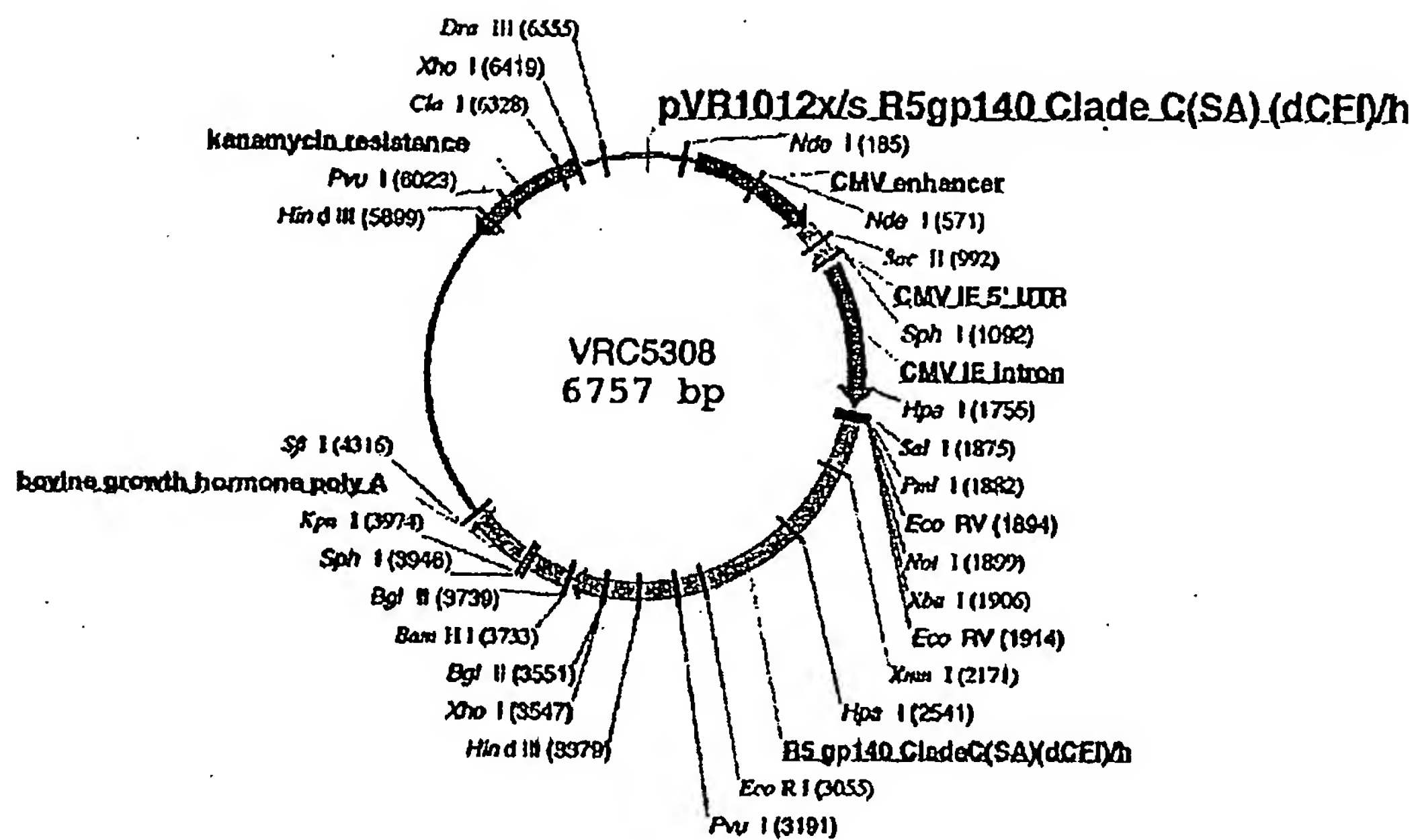


Figure 95

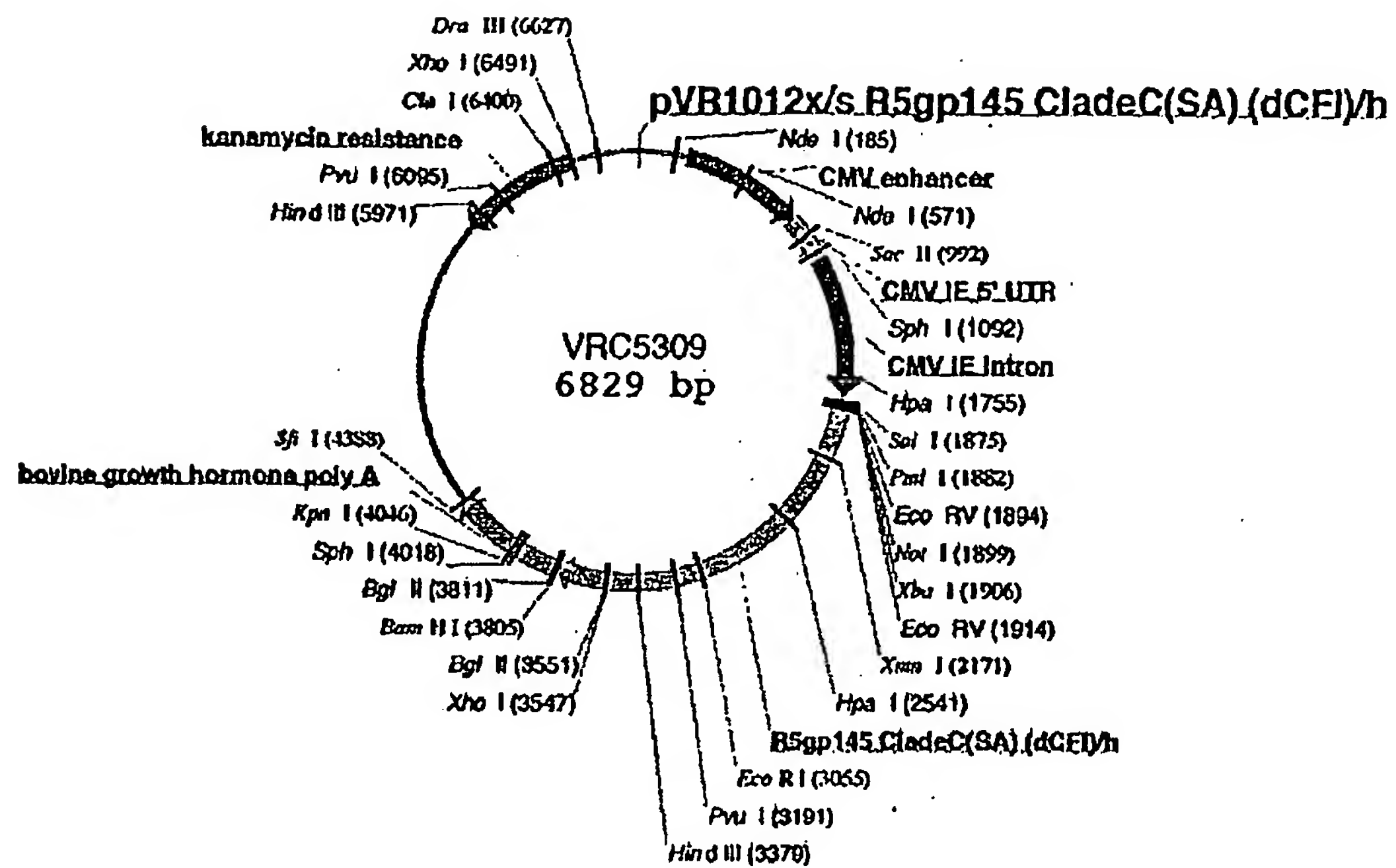


Figure 96

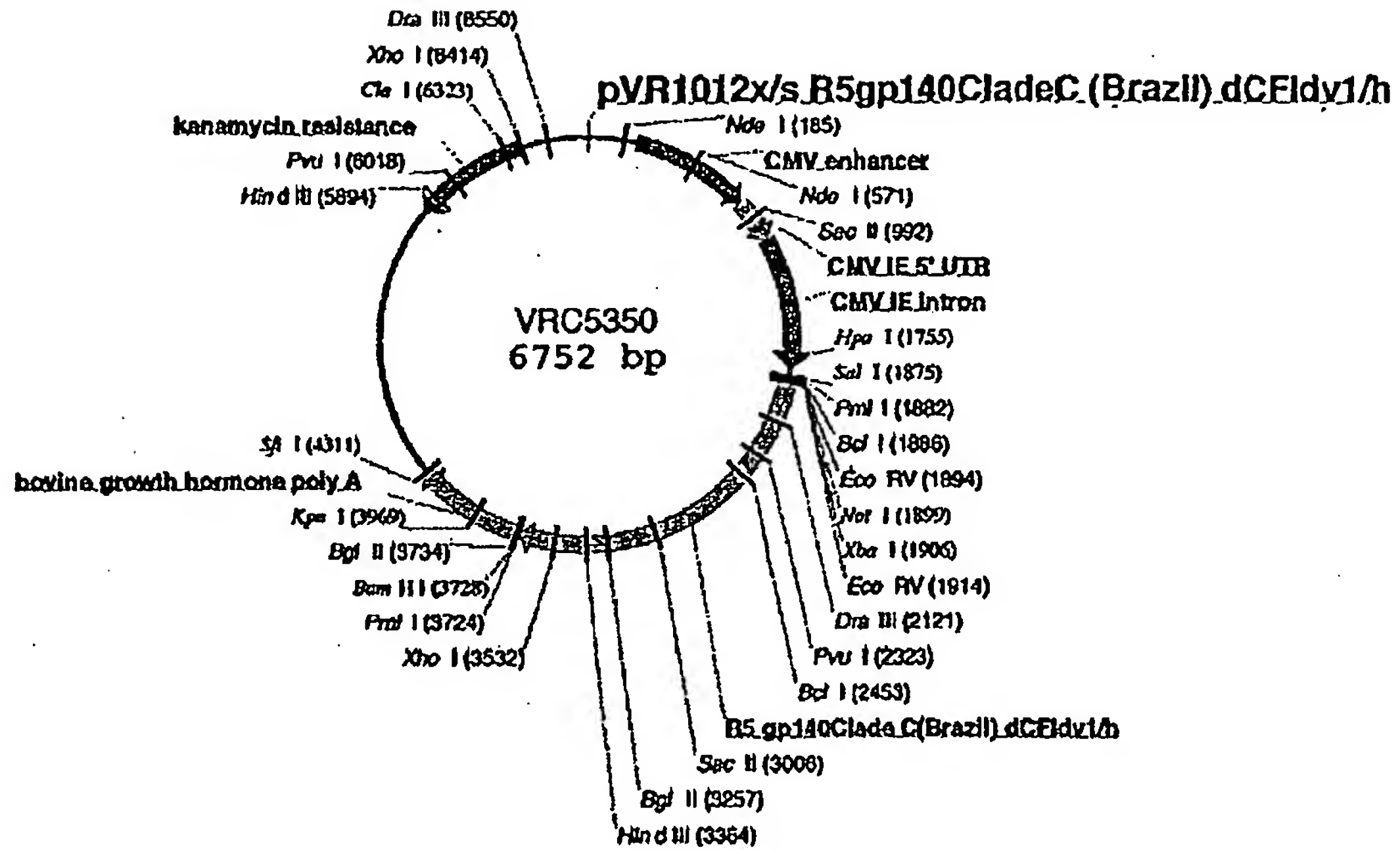


Figure 97



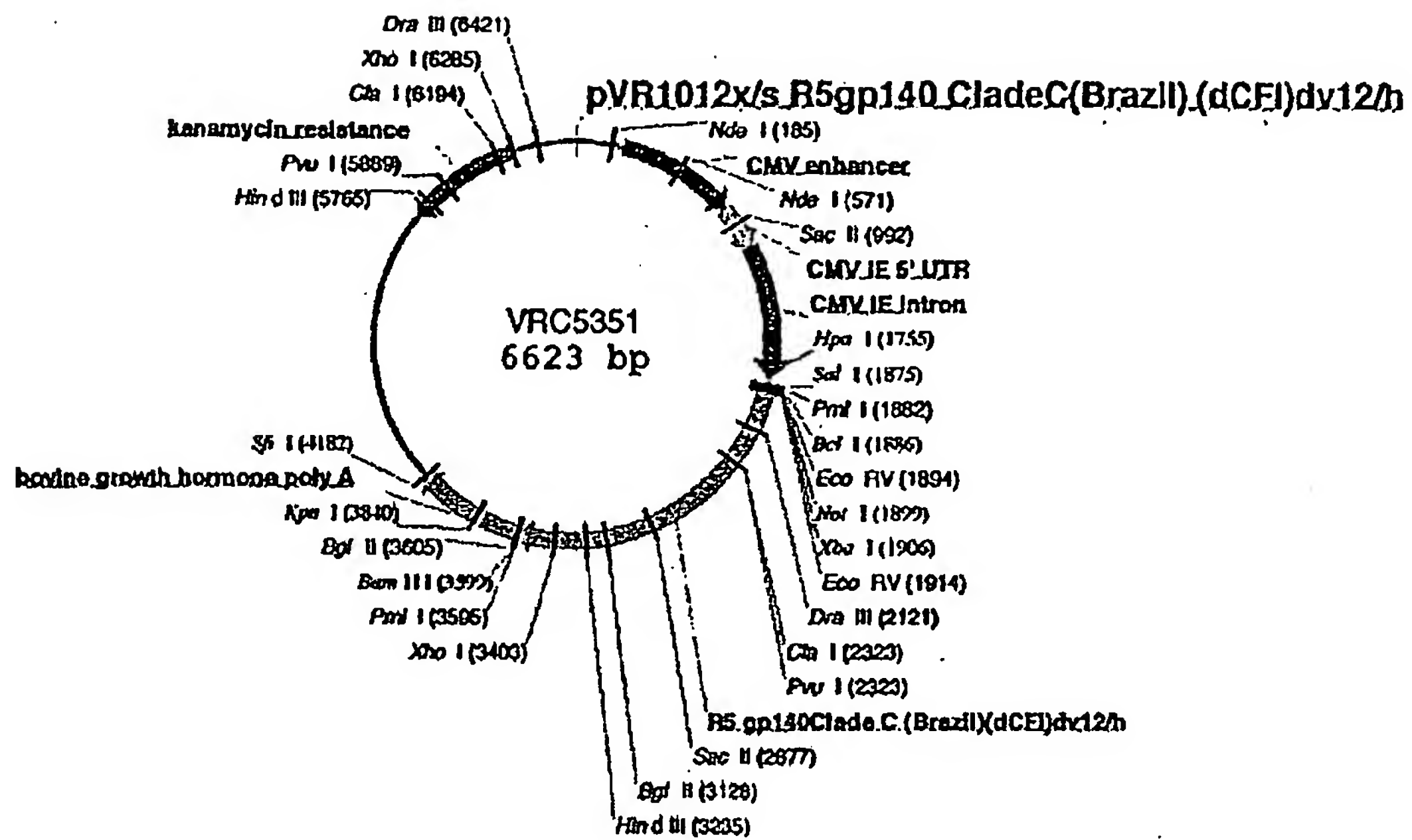


Figure 98

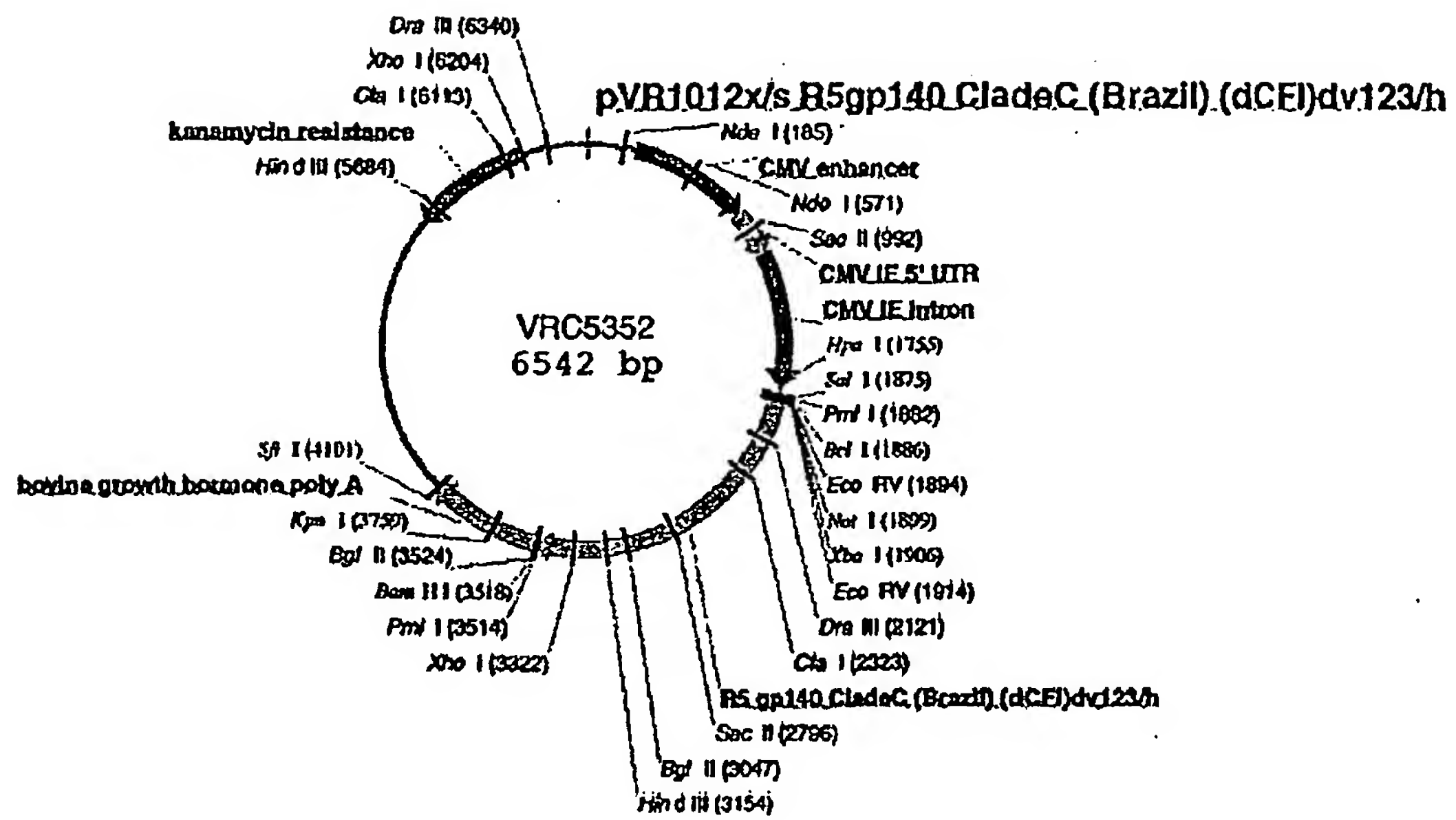


Figure 99

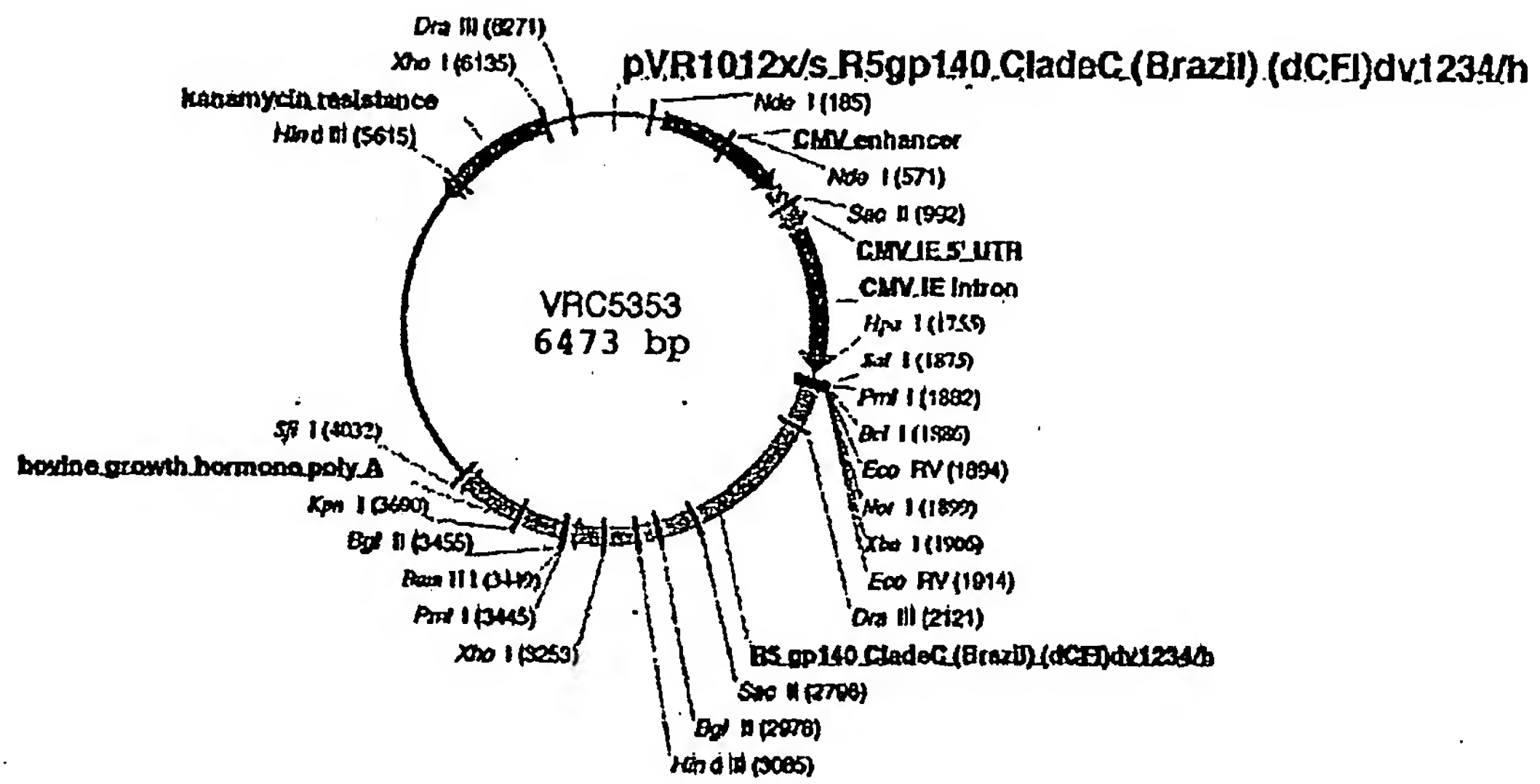


Figure 100

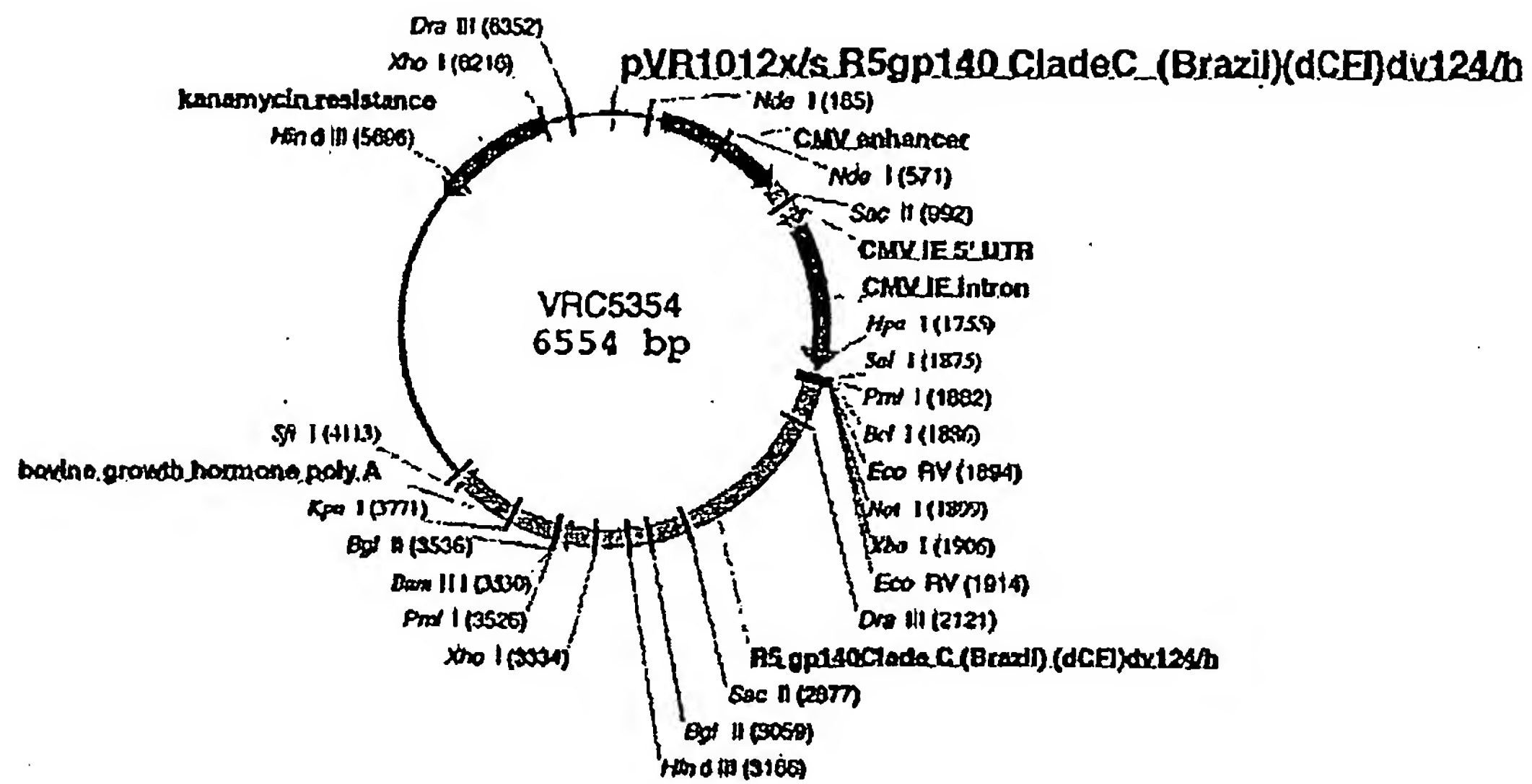


Figure 101

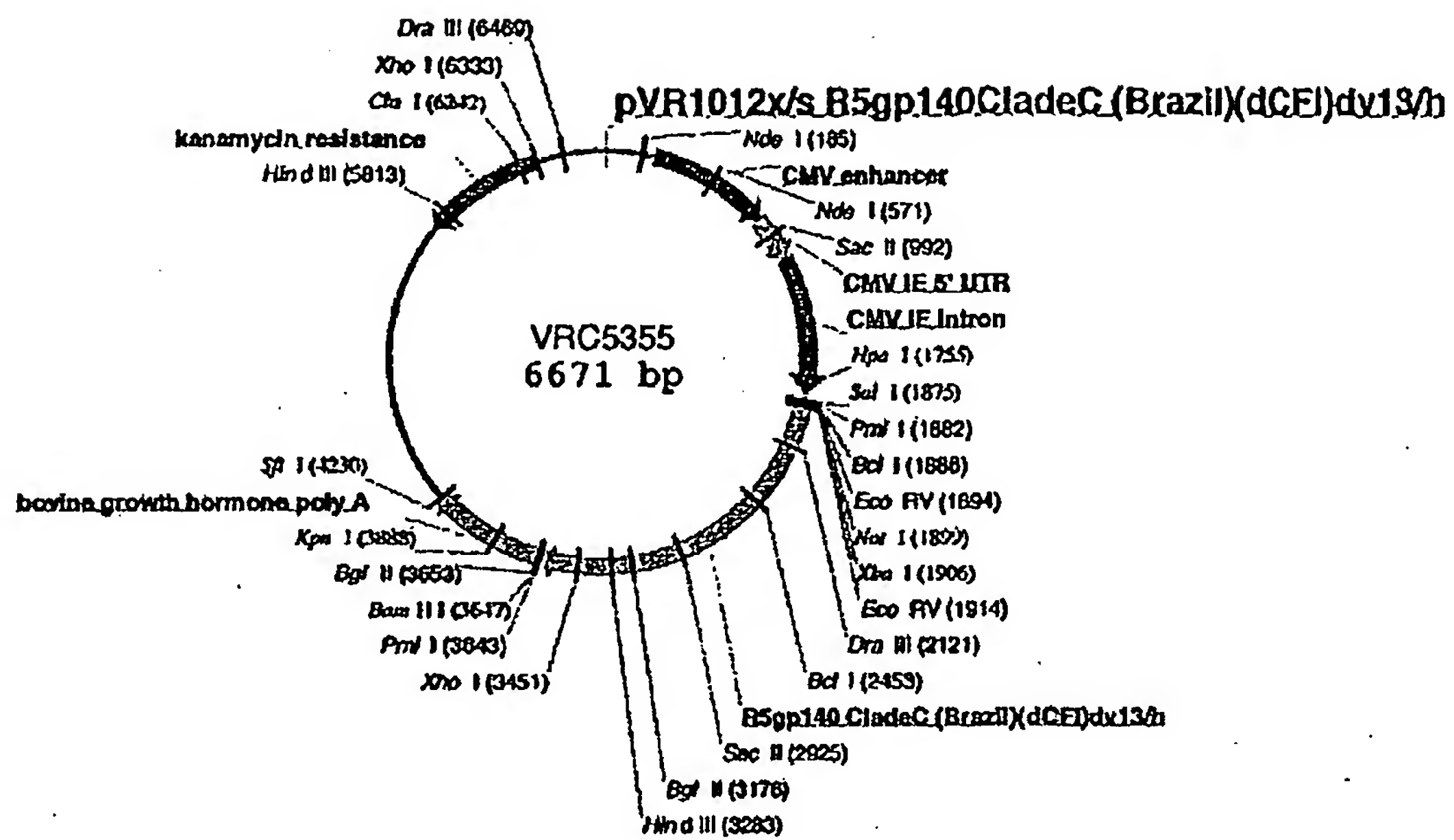


Figure 102

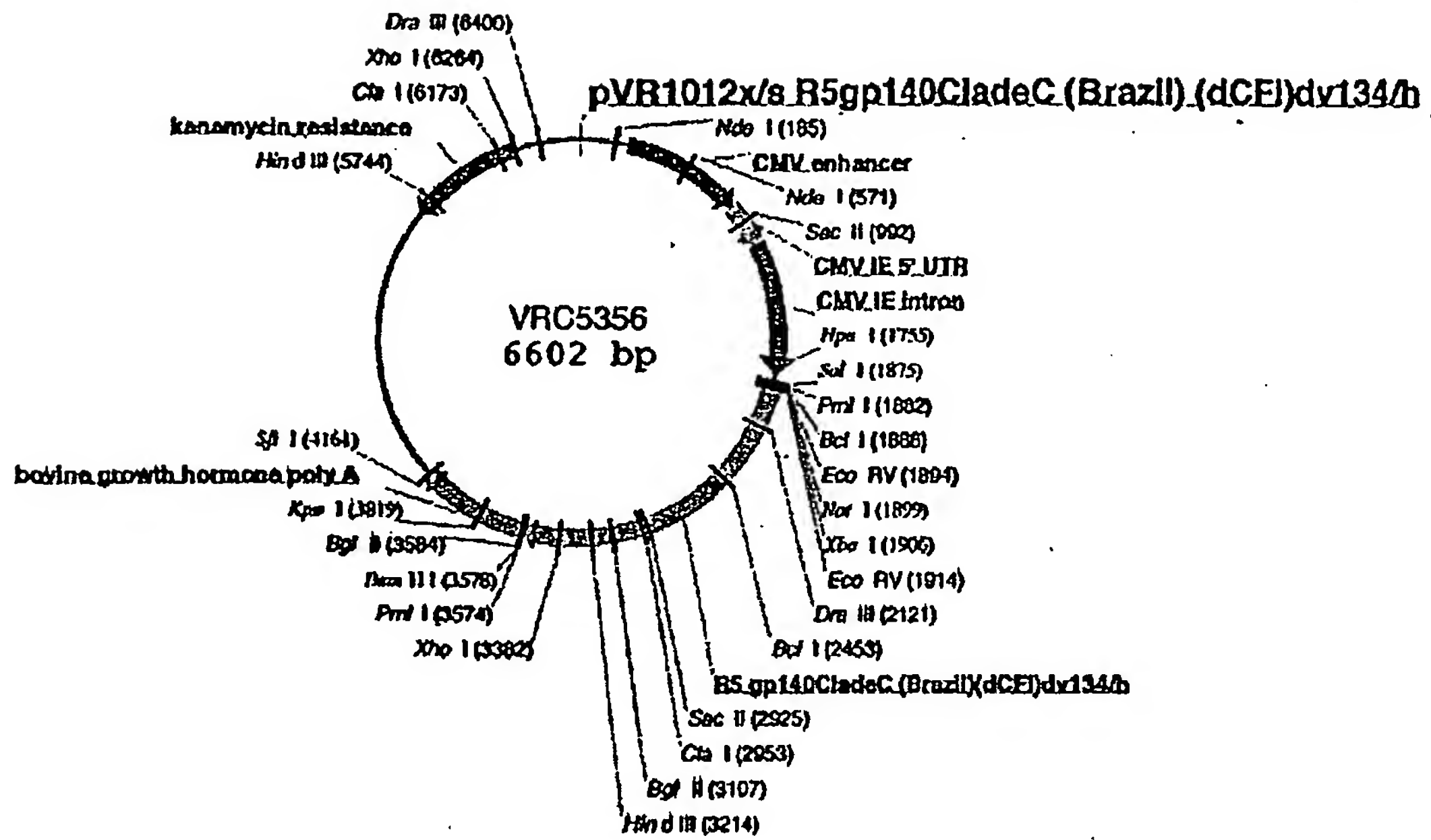


Figure 103

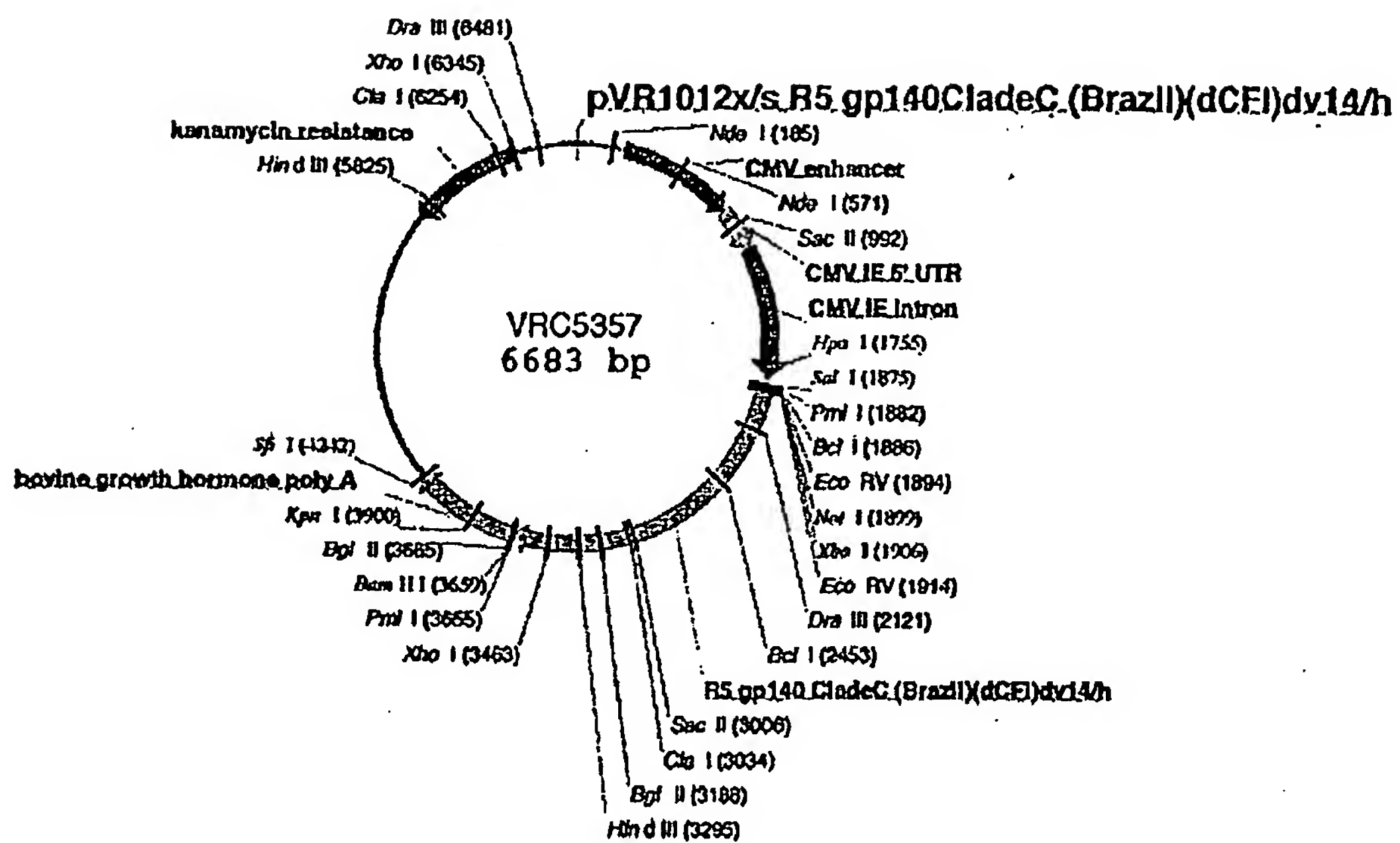


Figure 104



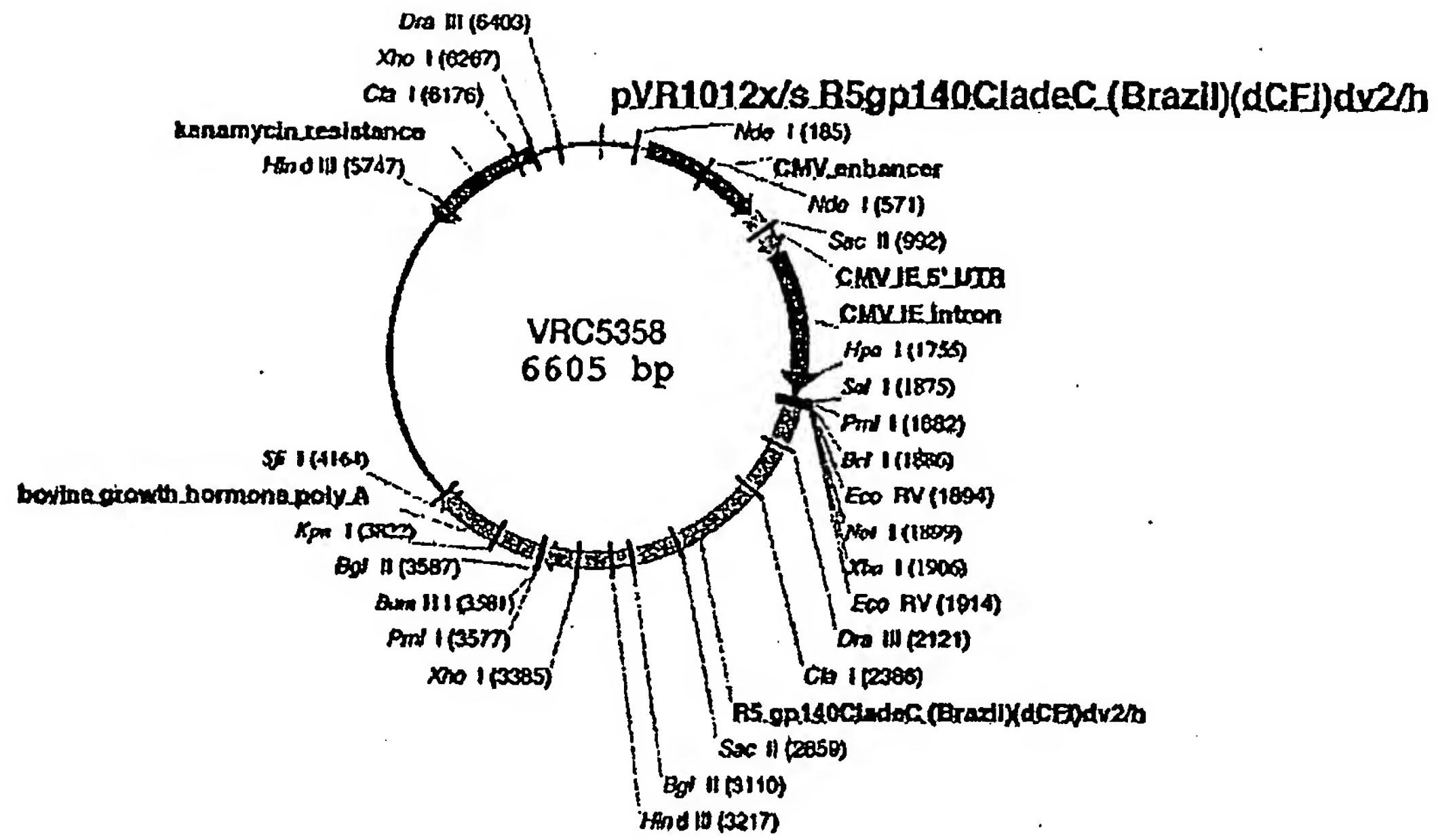


Figure 105

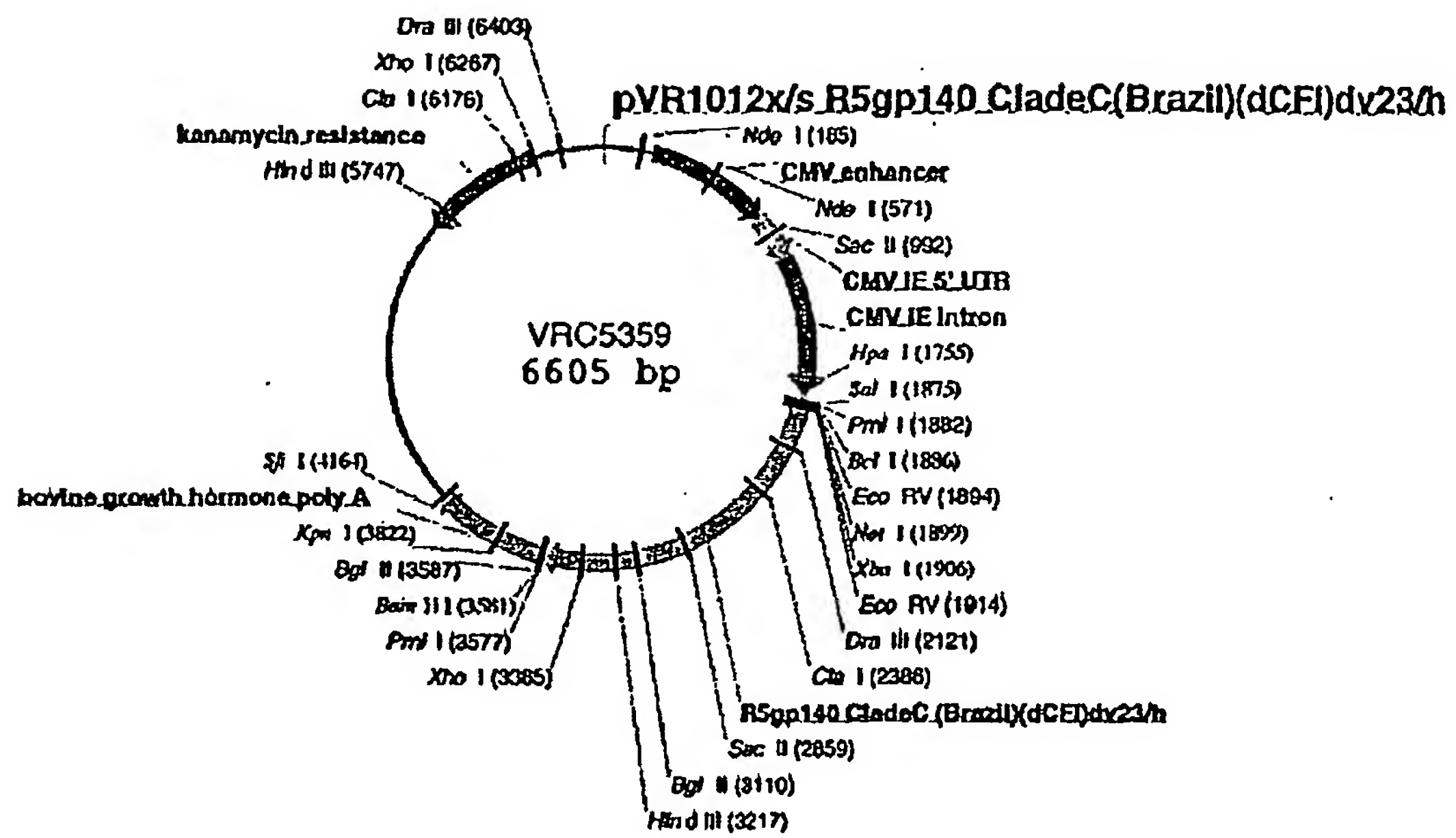


Figure 106

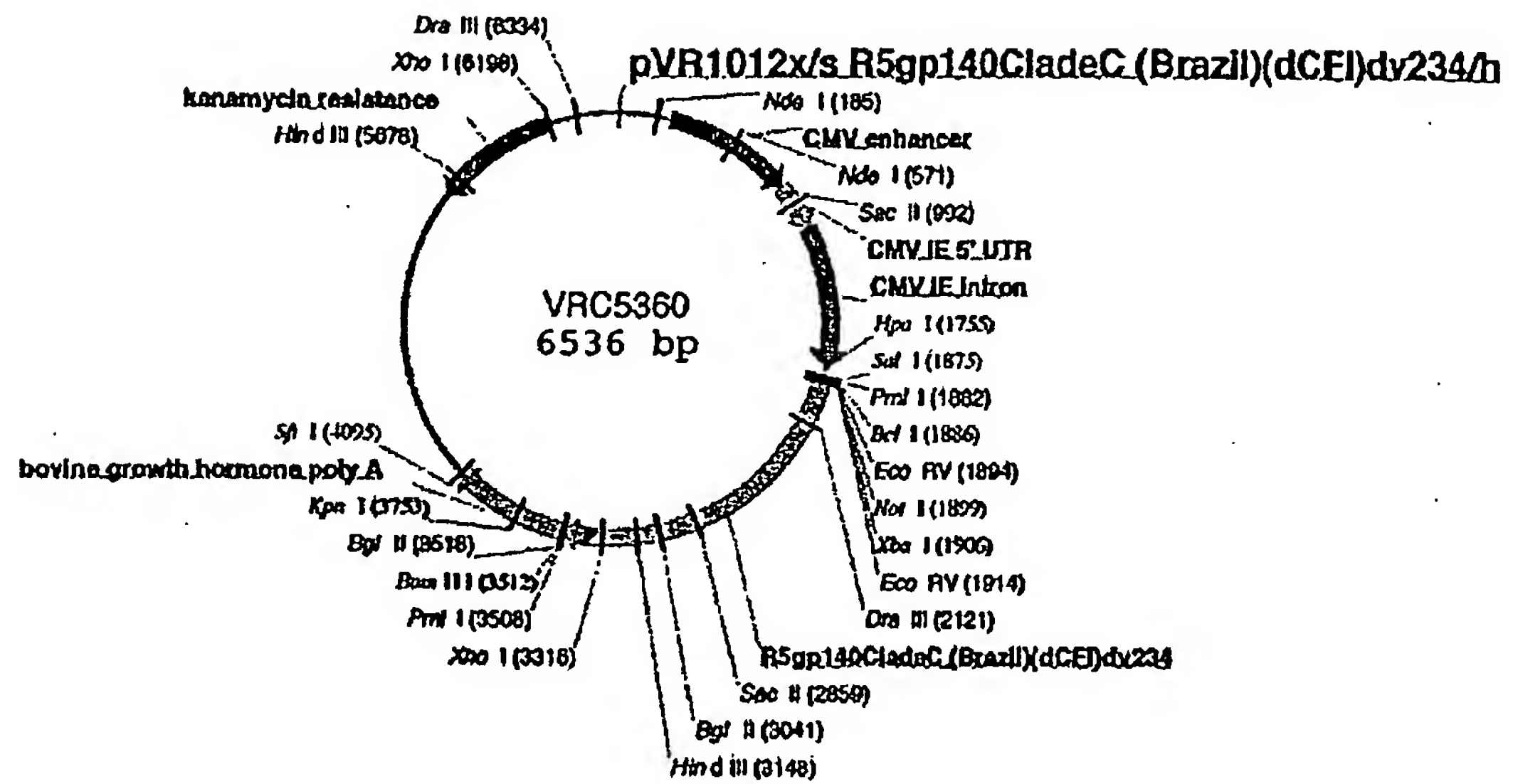


Figure 107

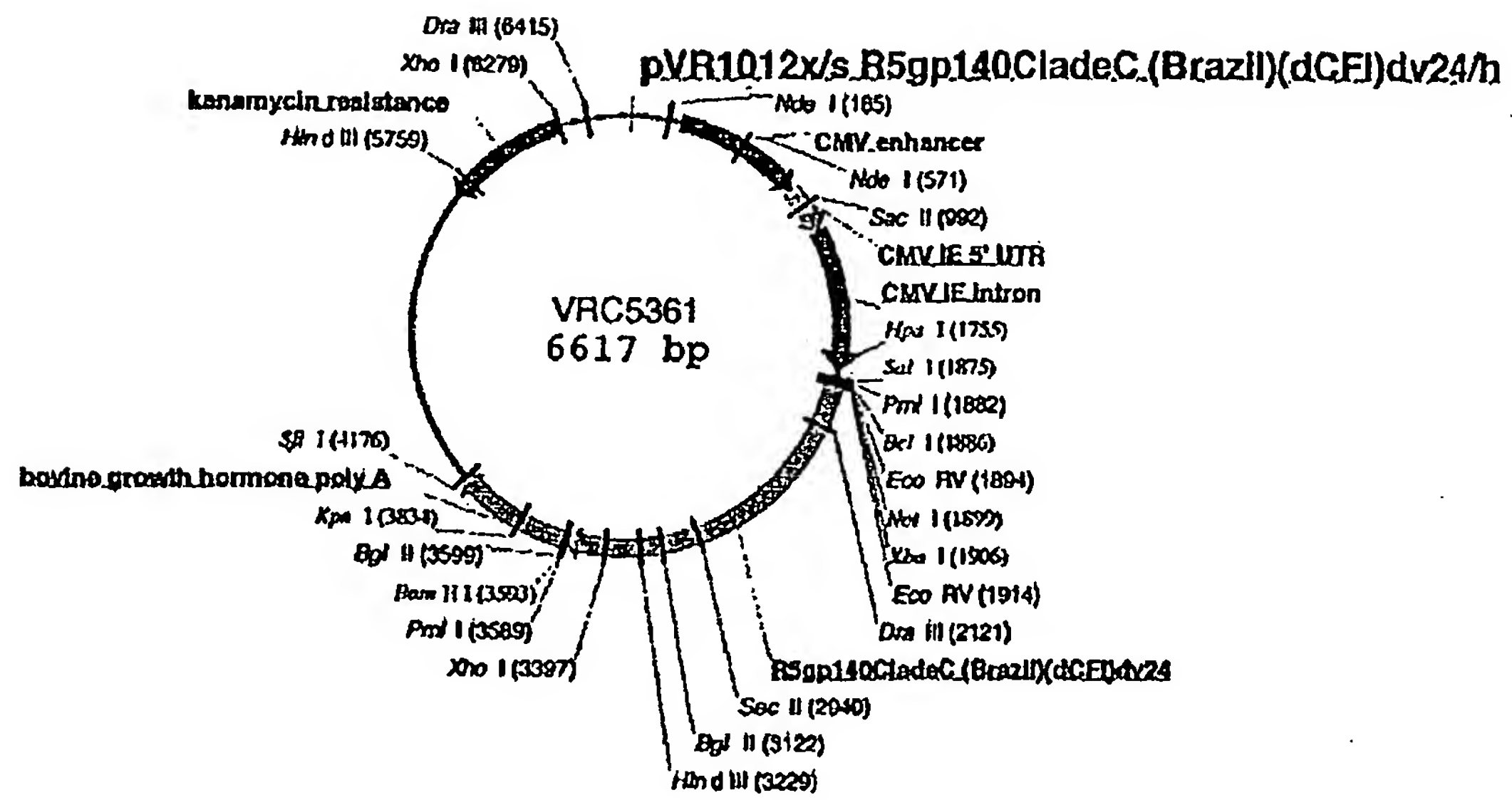


Figure 108

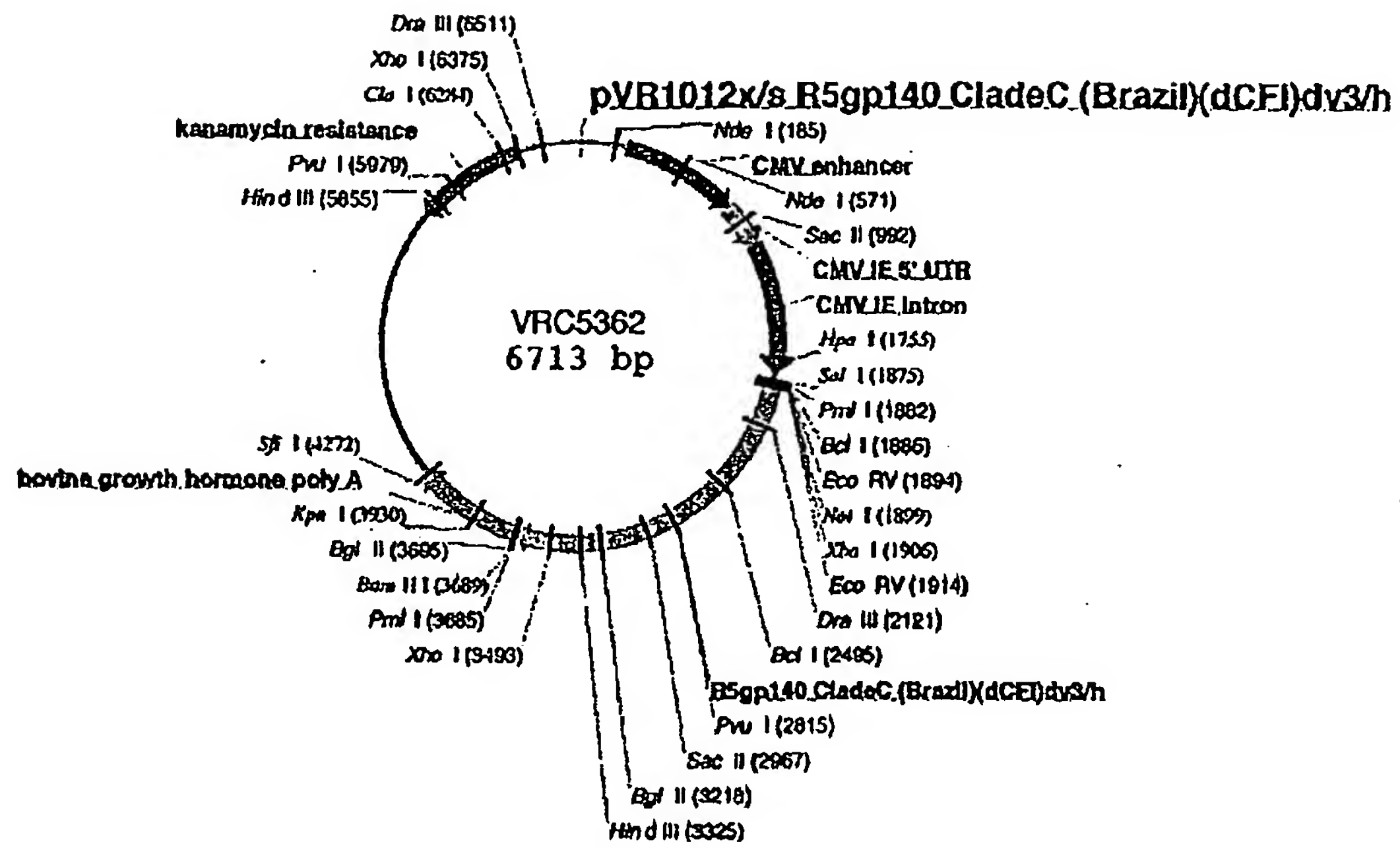


Figure 109

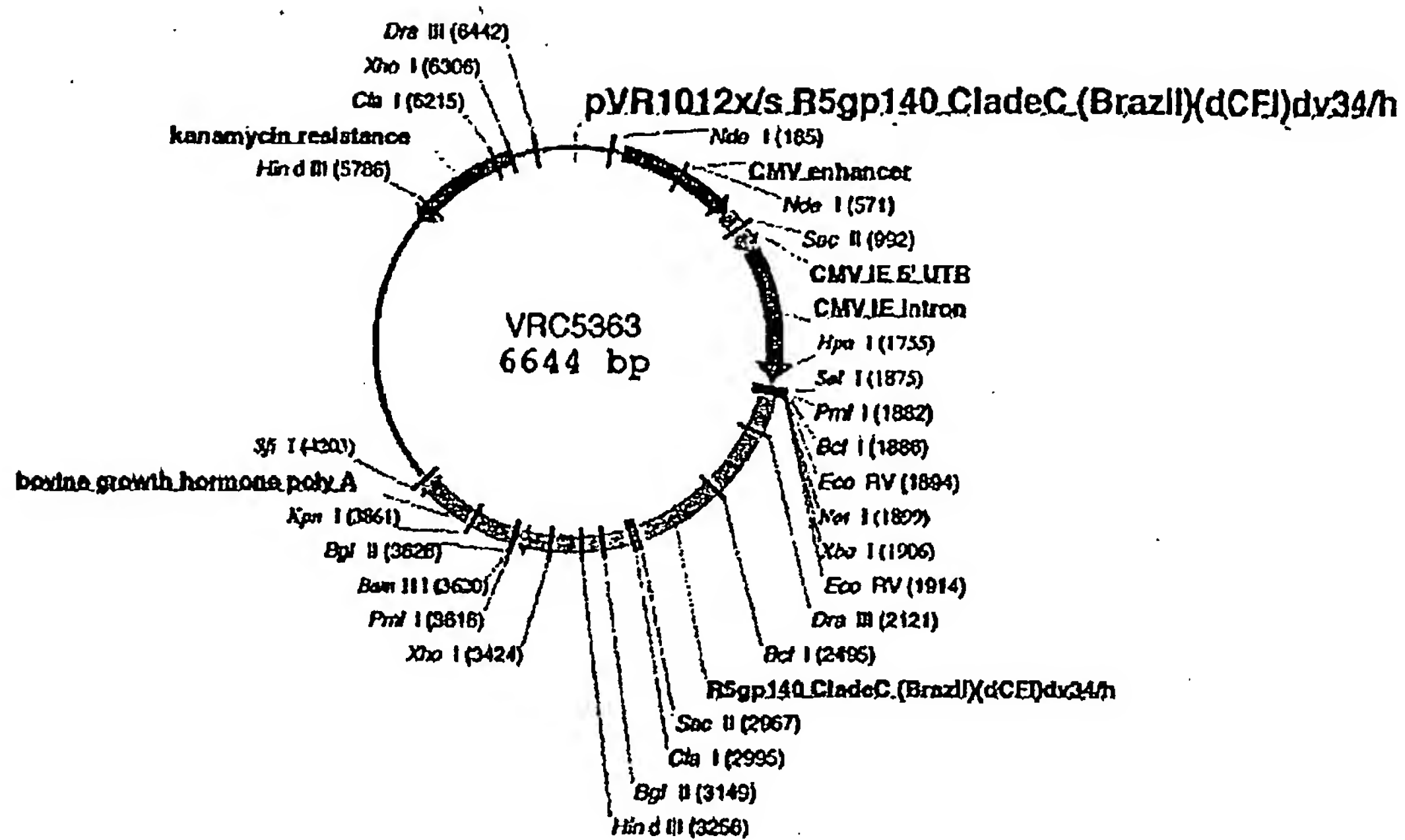


Figure 110

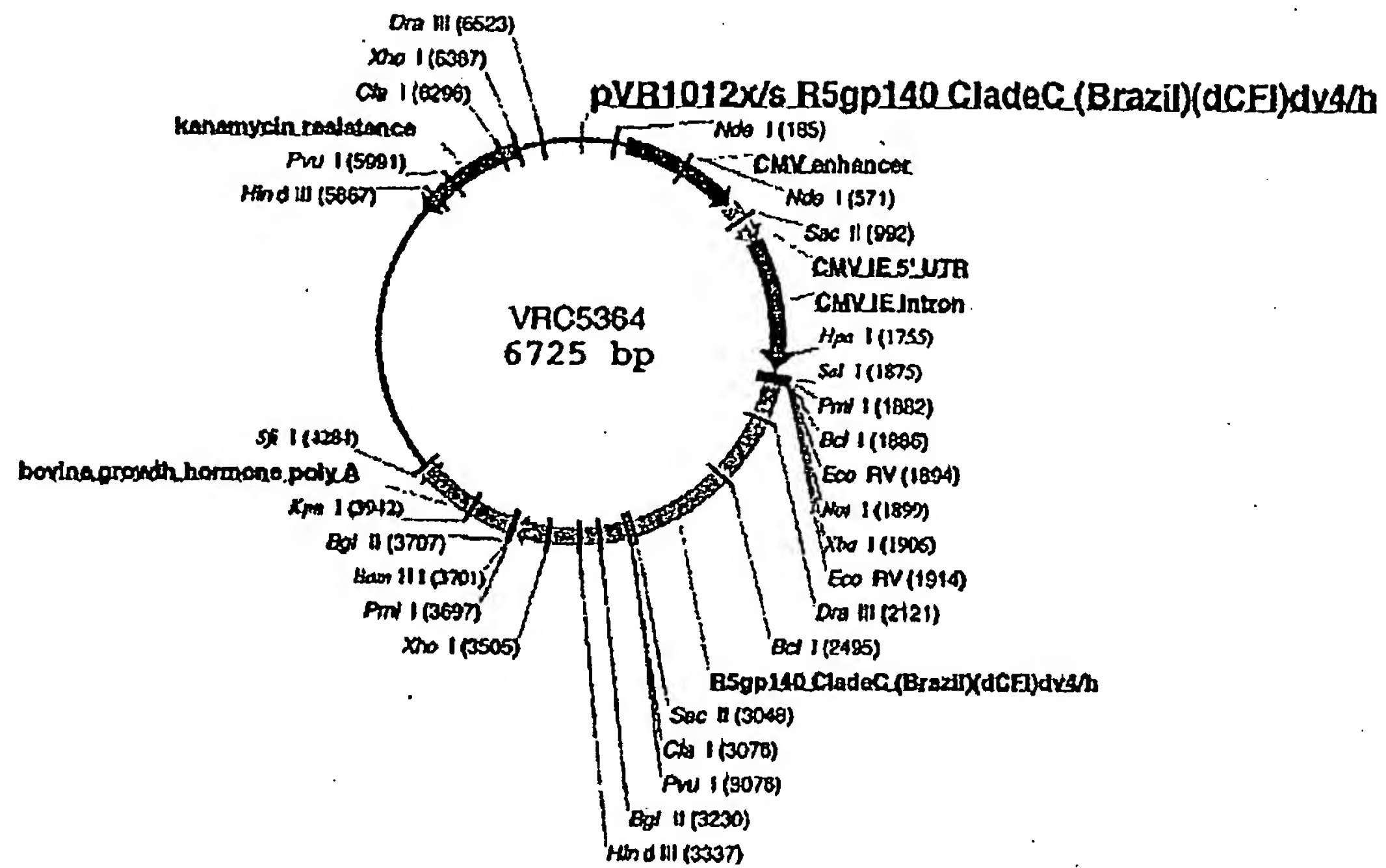


Figure 111



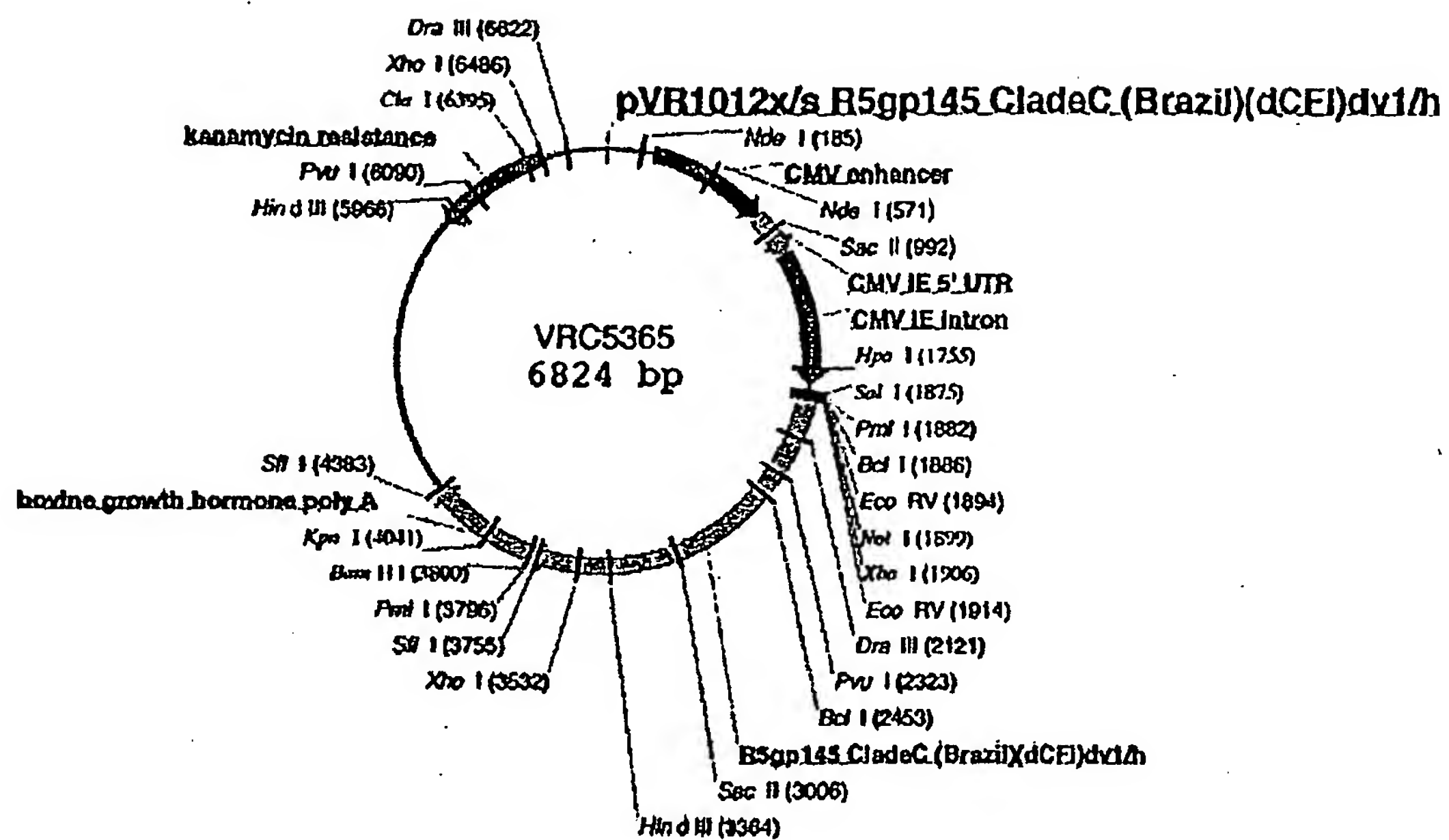


Figure 112

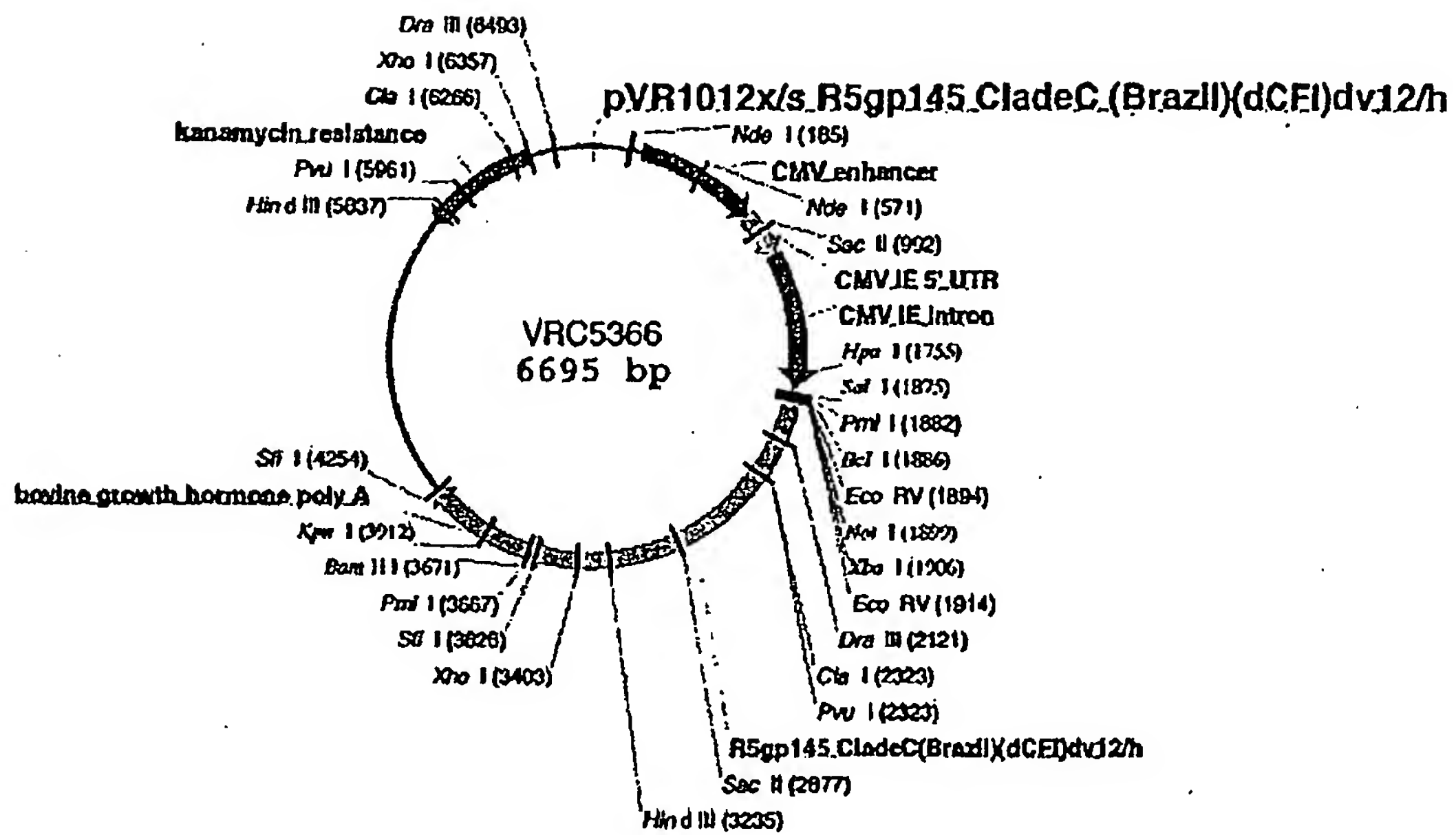


Figure 113

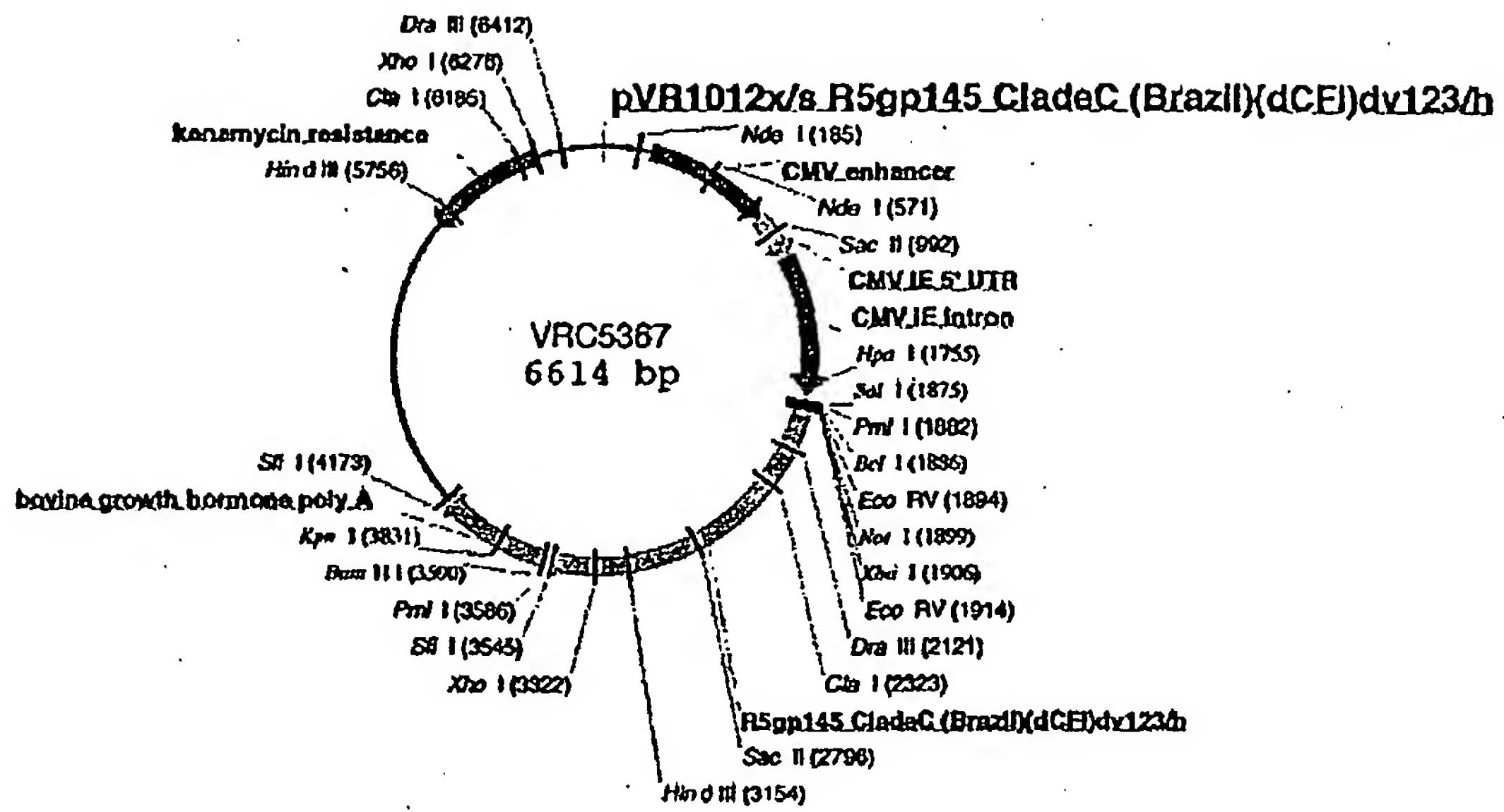


Figure 114

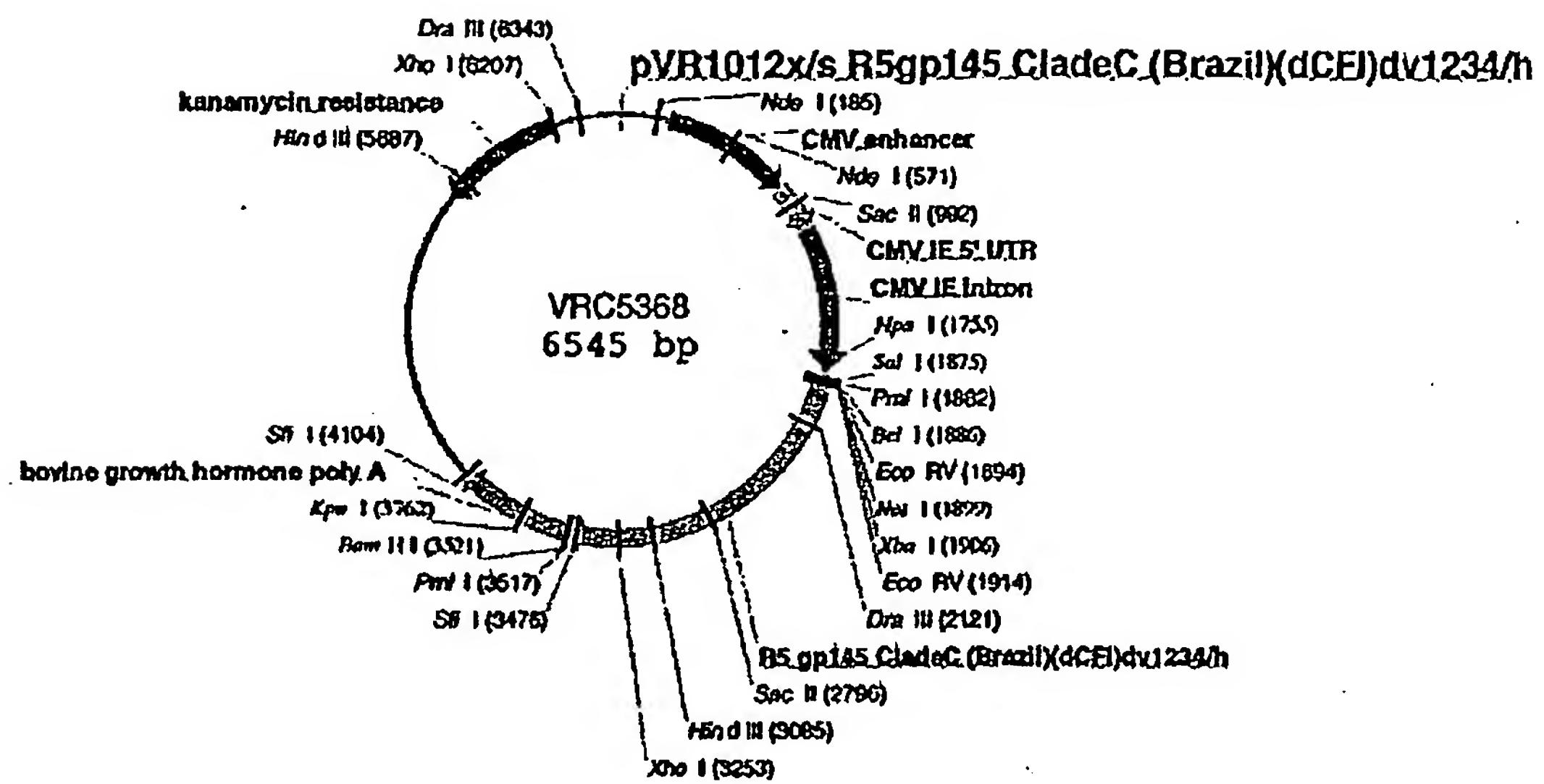


Figure 115

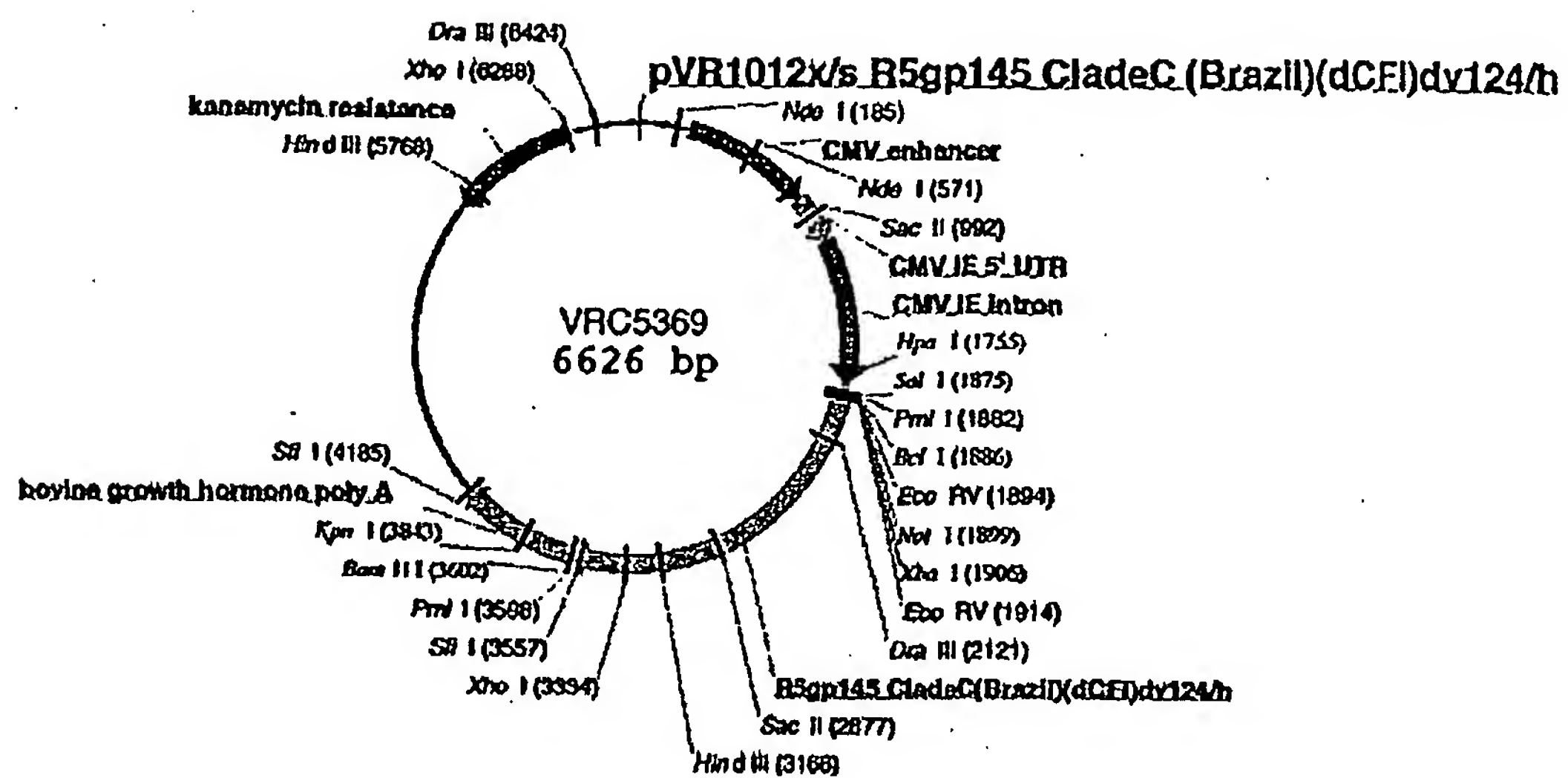


Figure 116

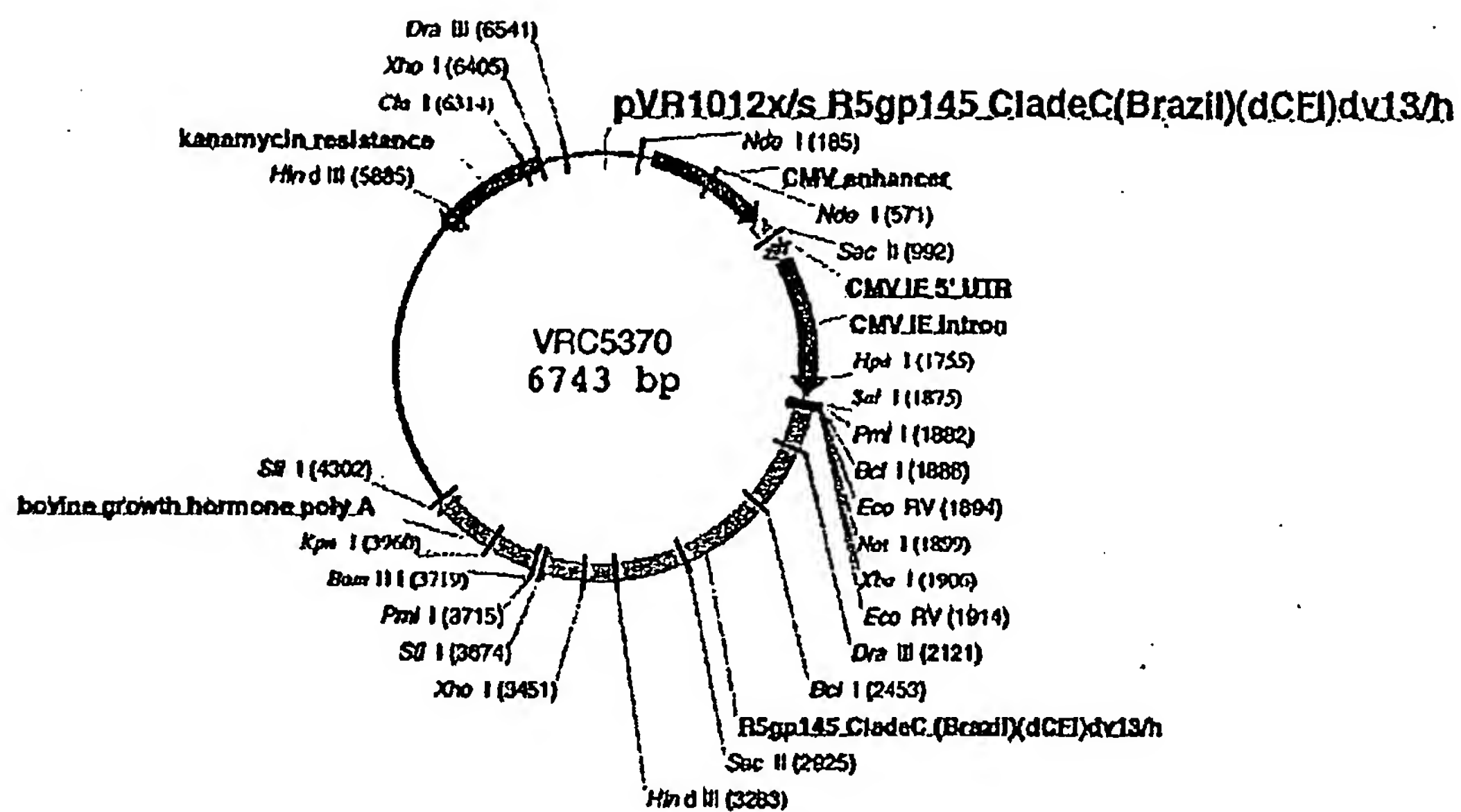


Figure 117

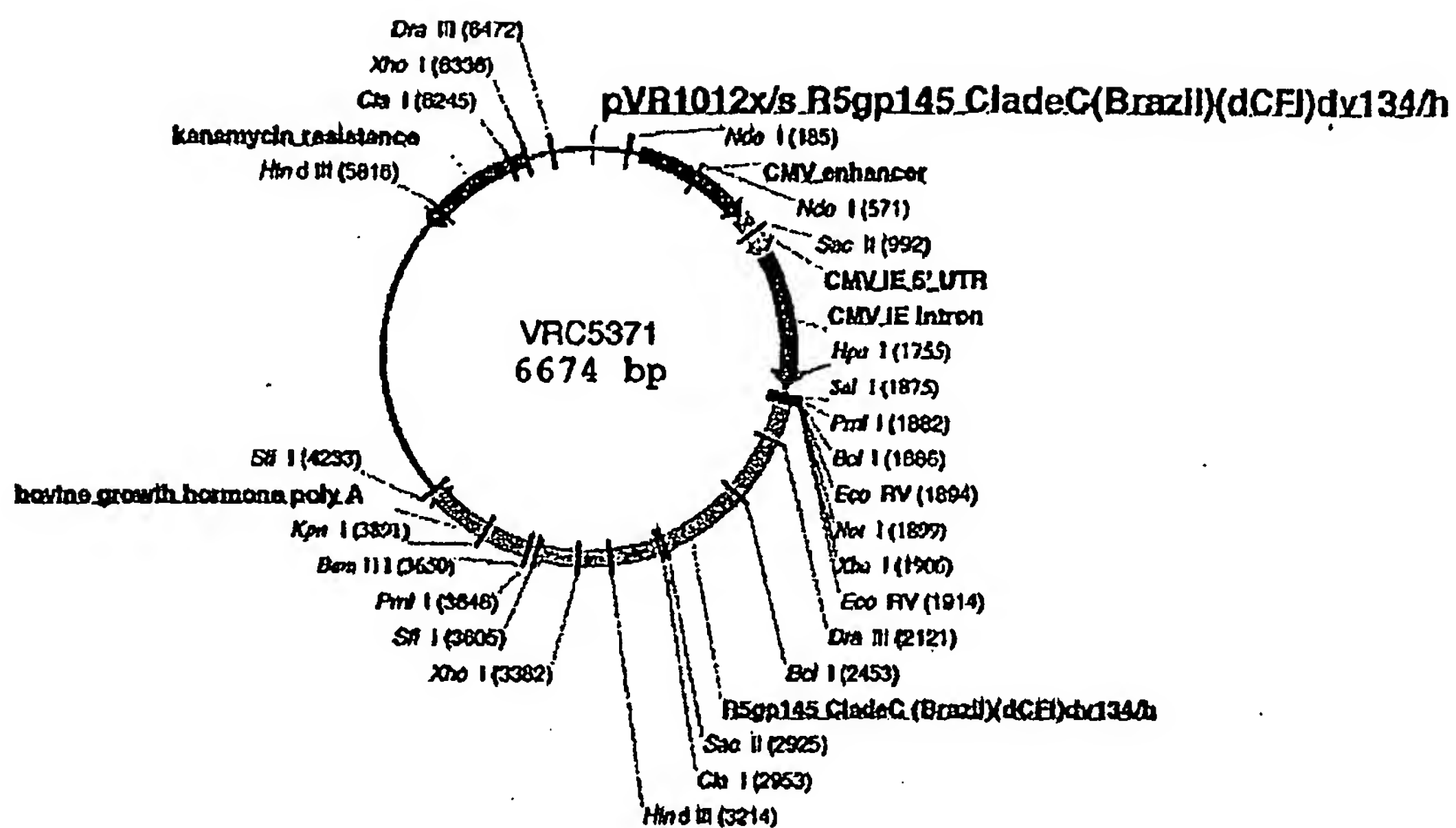


Figure 118



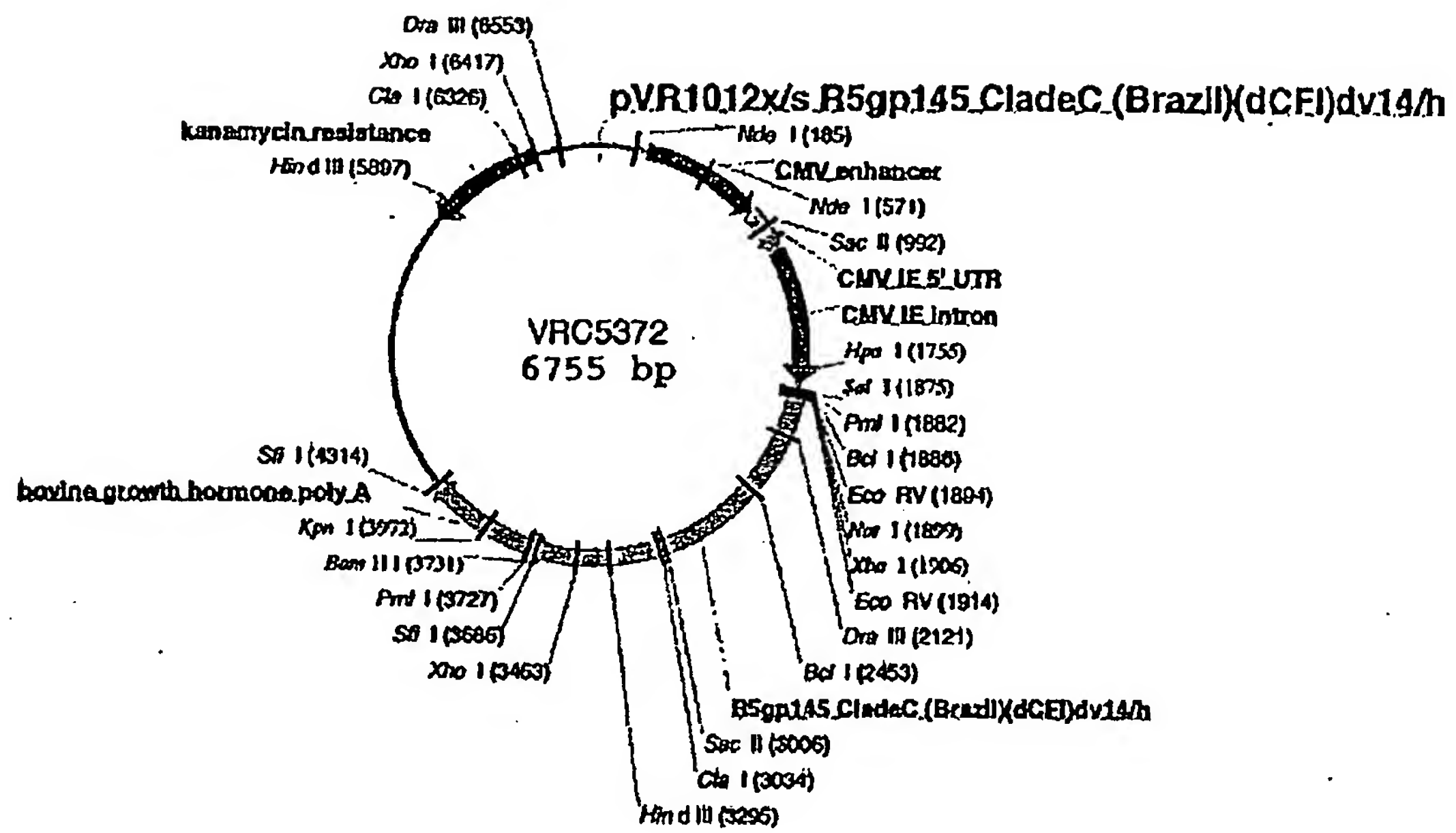


Figure 119

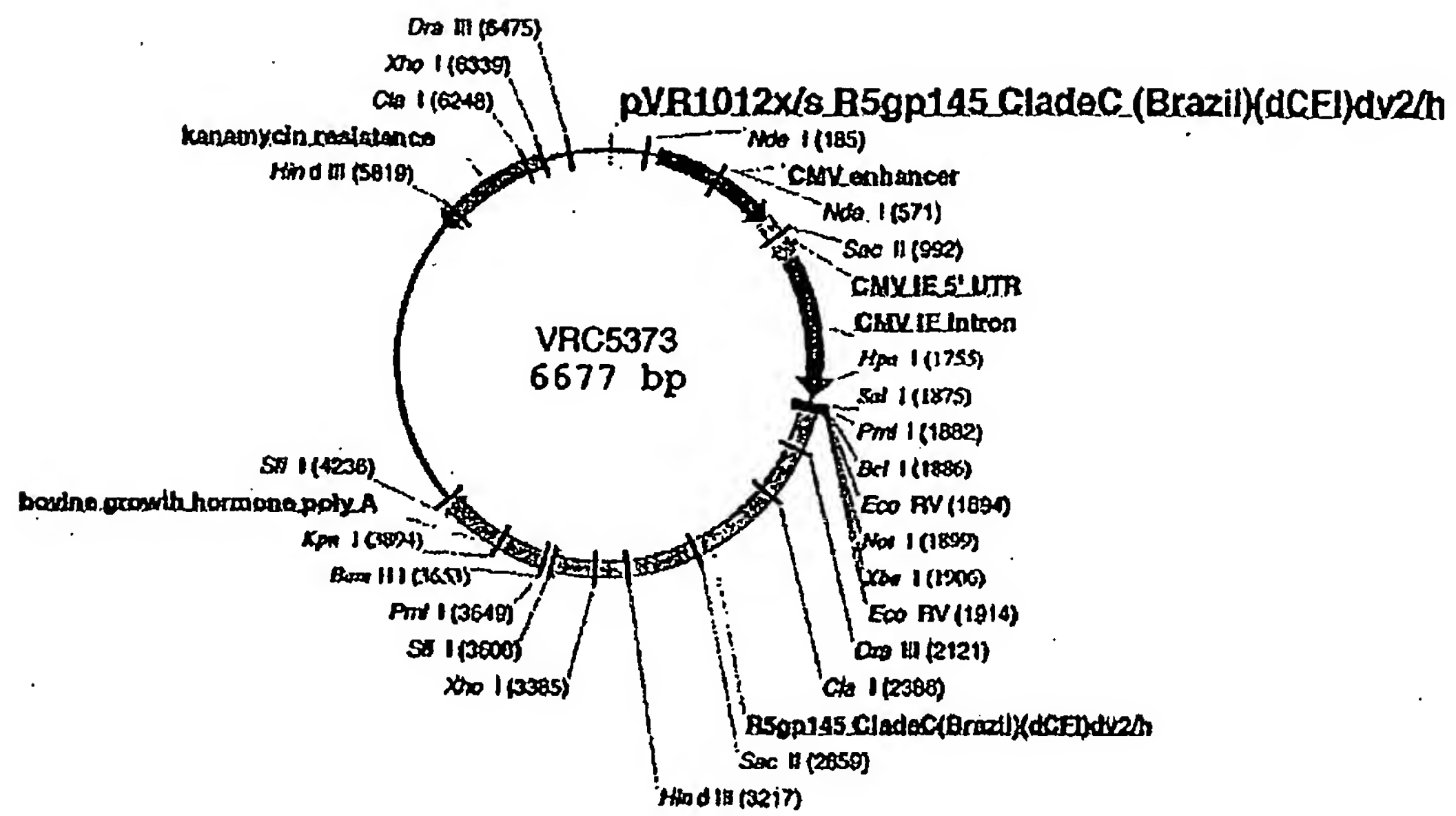


Figure 120

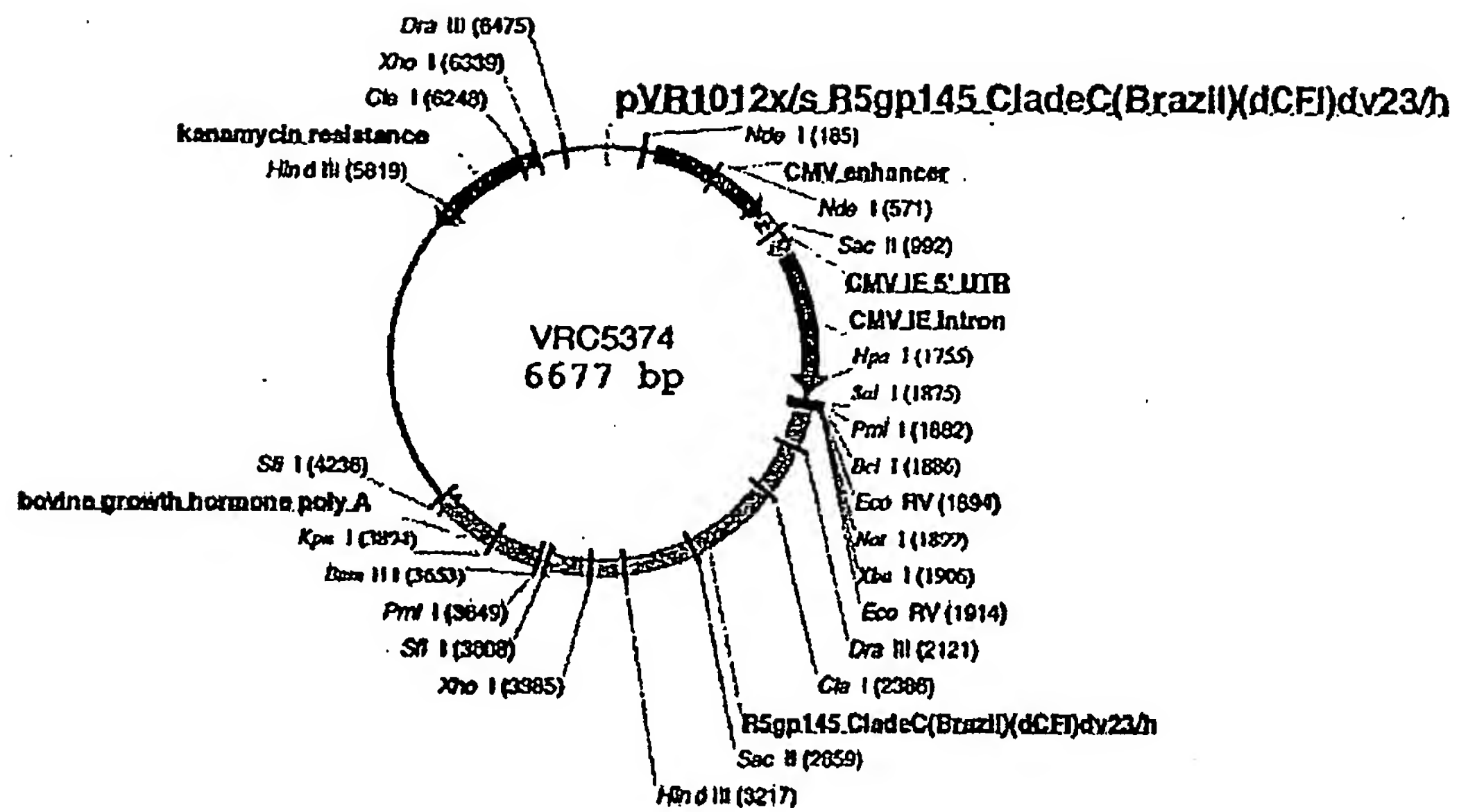


Figure 121

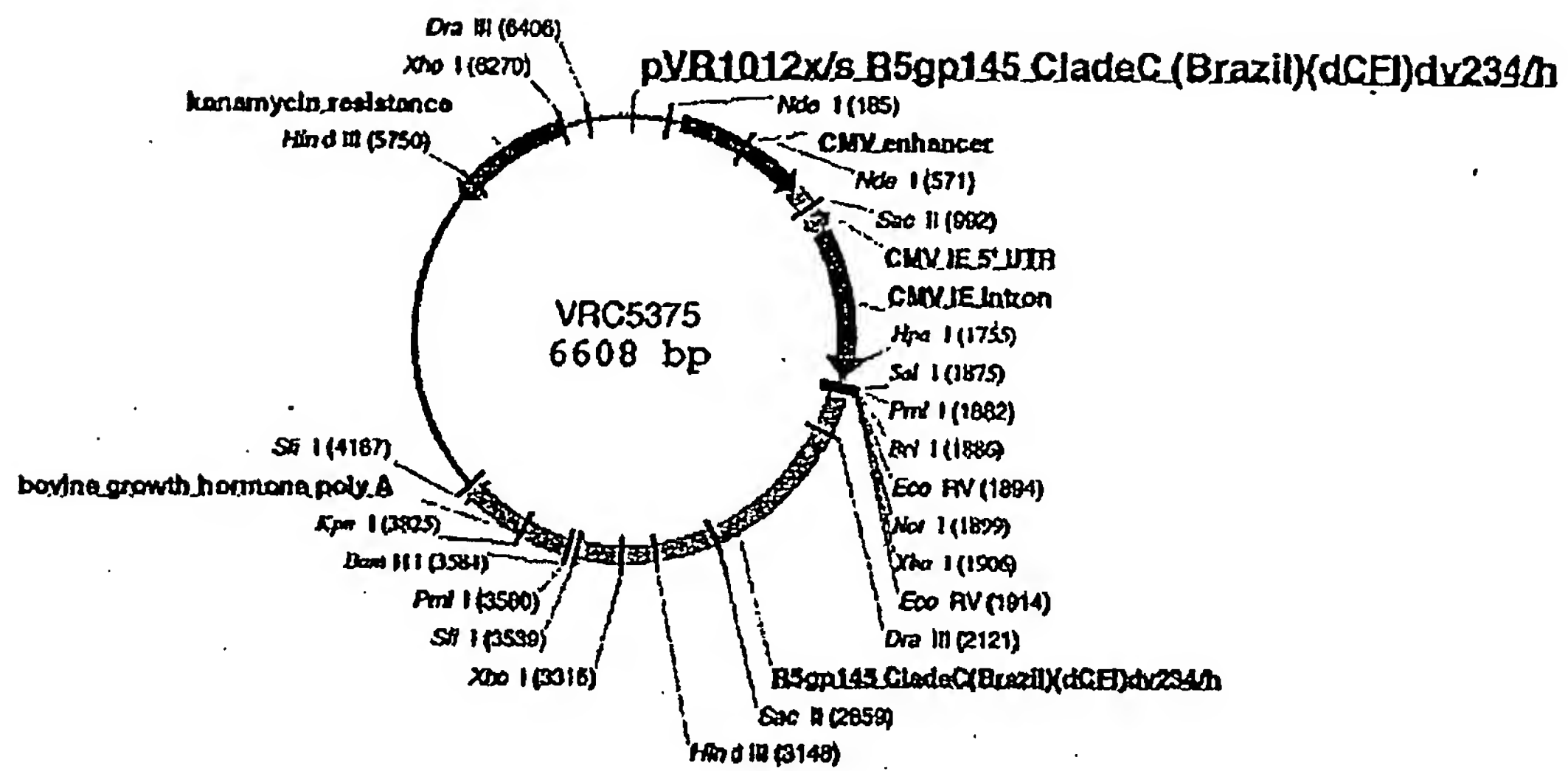


Figure 122

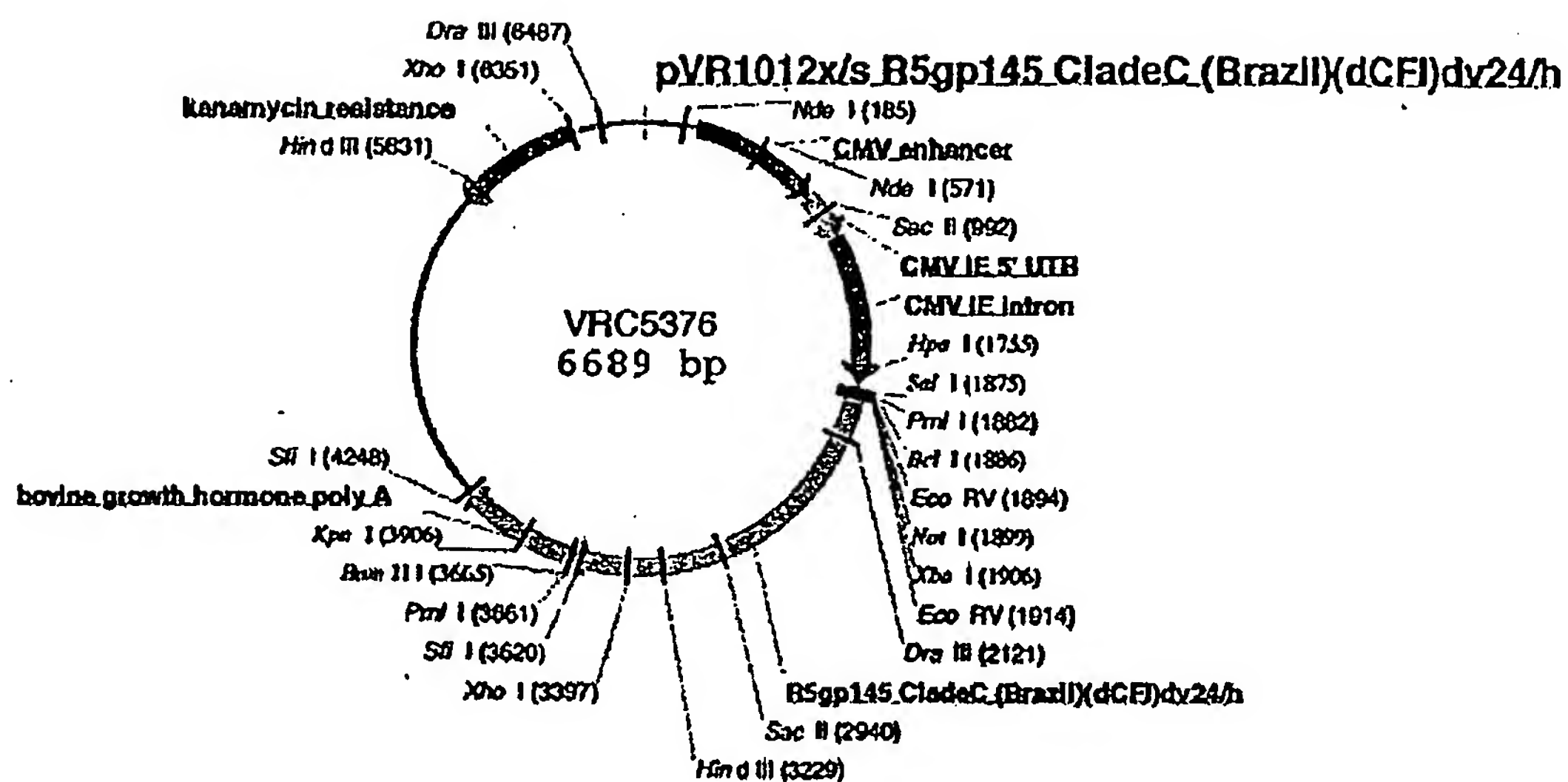


Figure 123

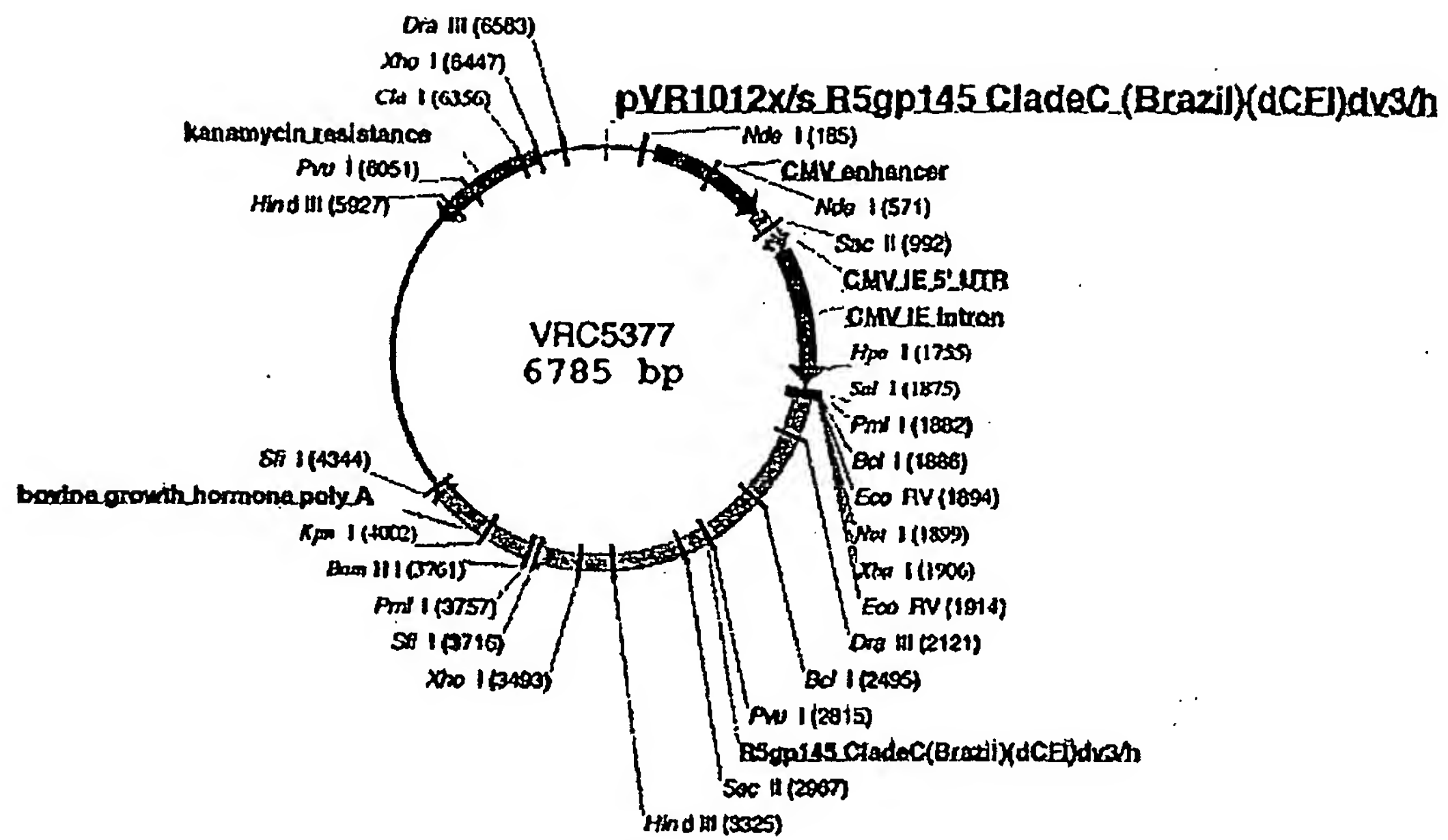


Figure 124

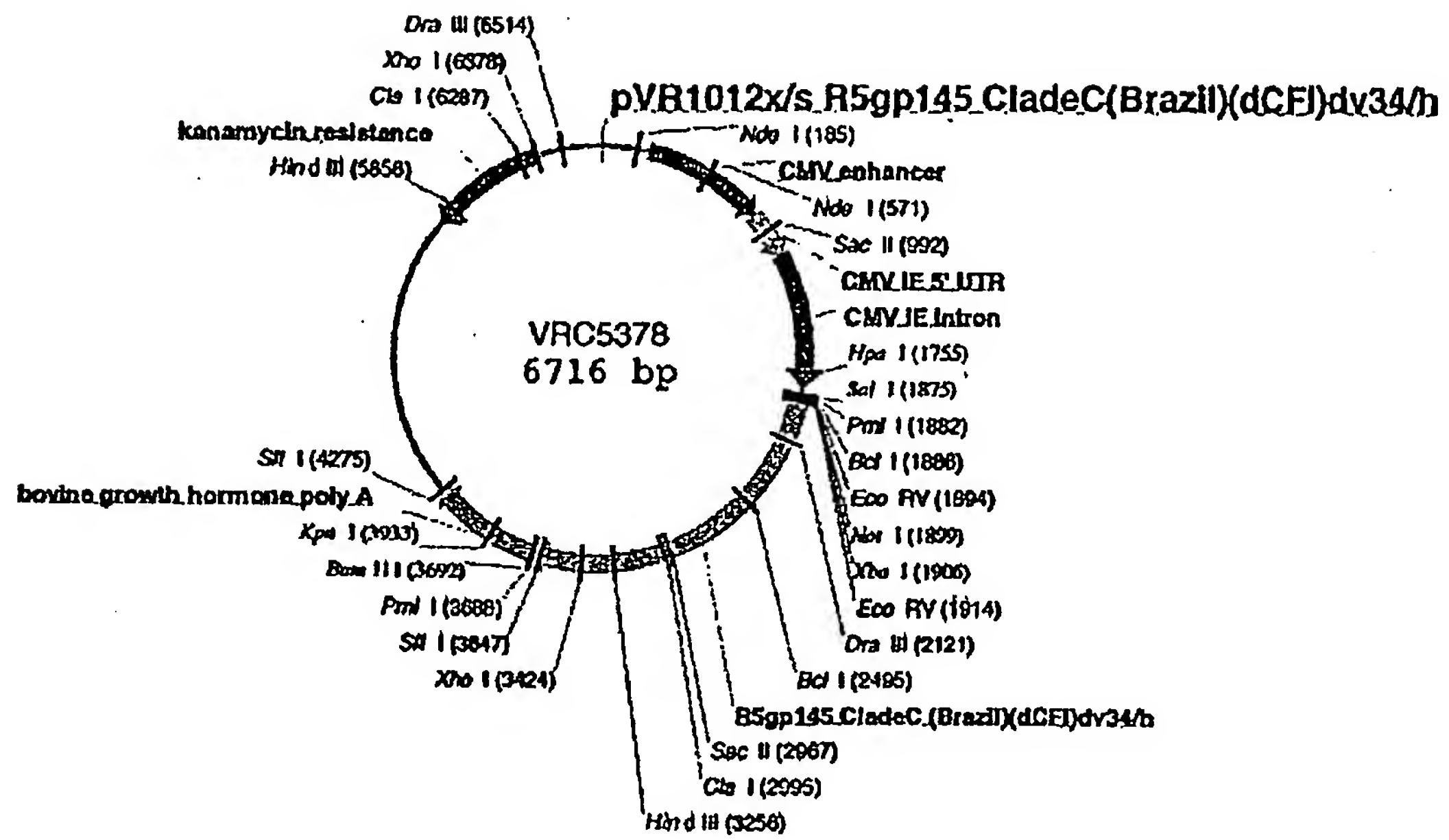


Figure 125



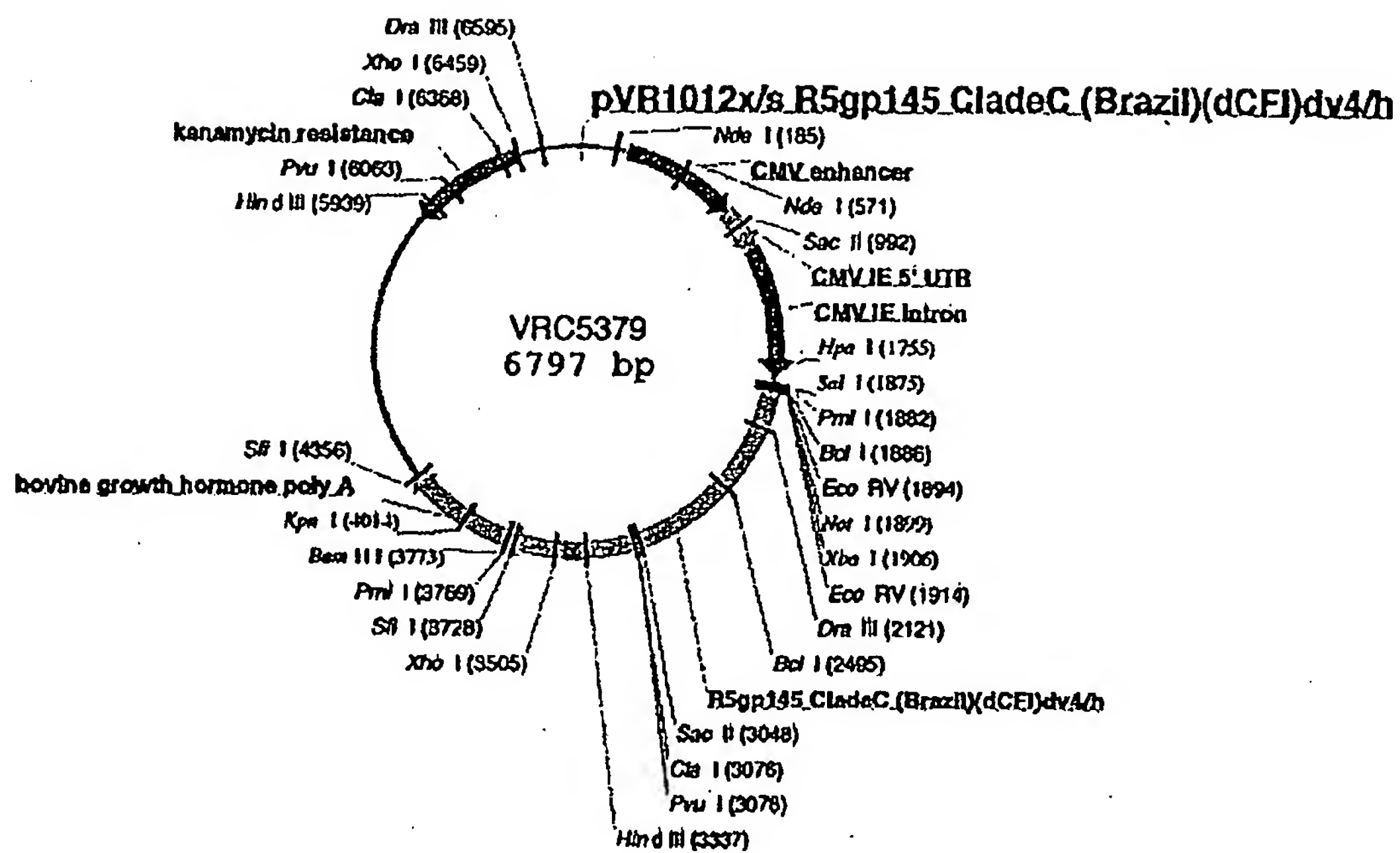


Figure 126

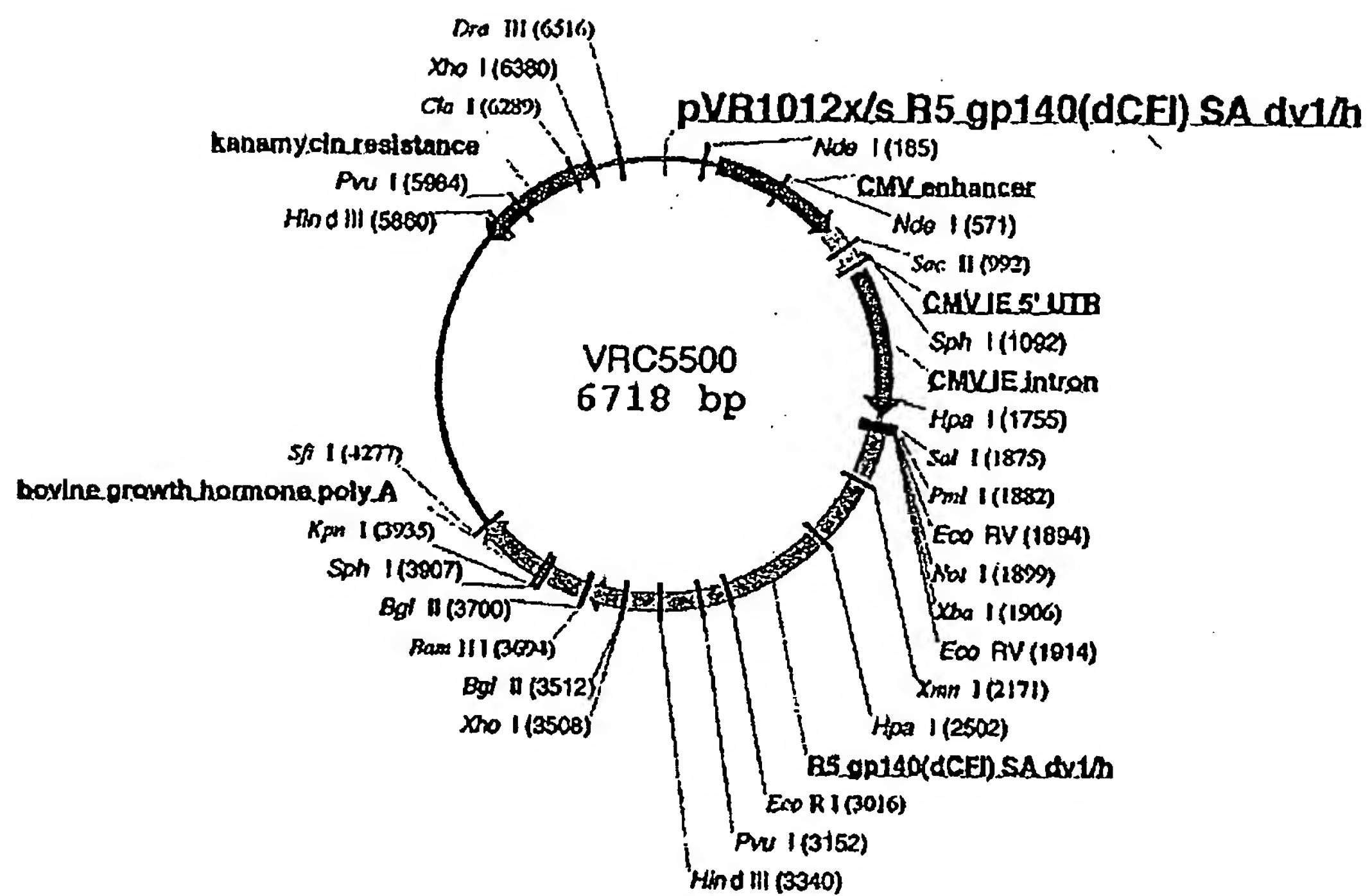


Figure 127

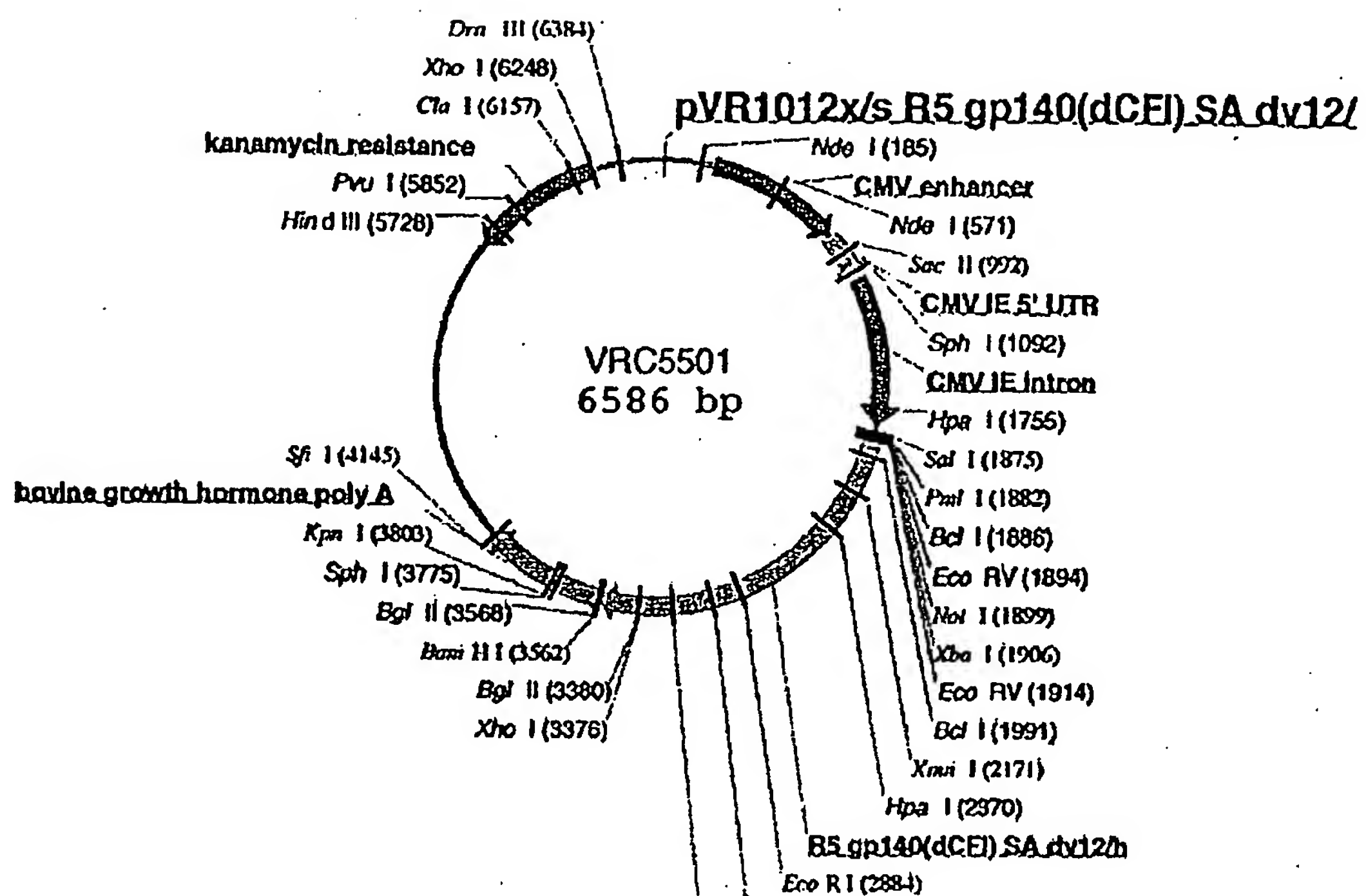


Figure 128

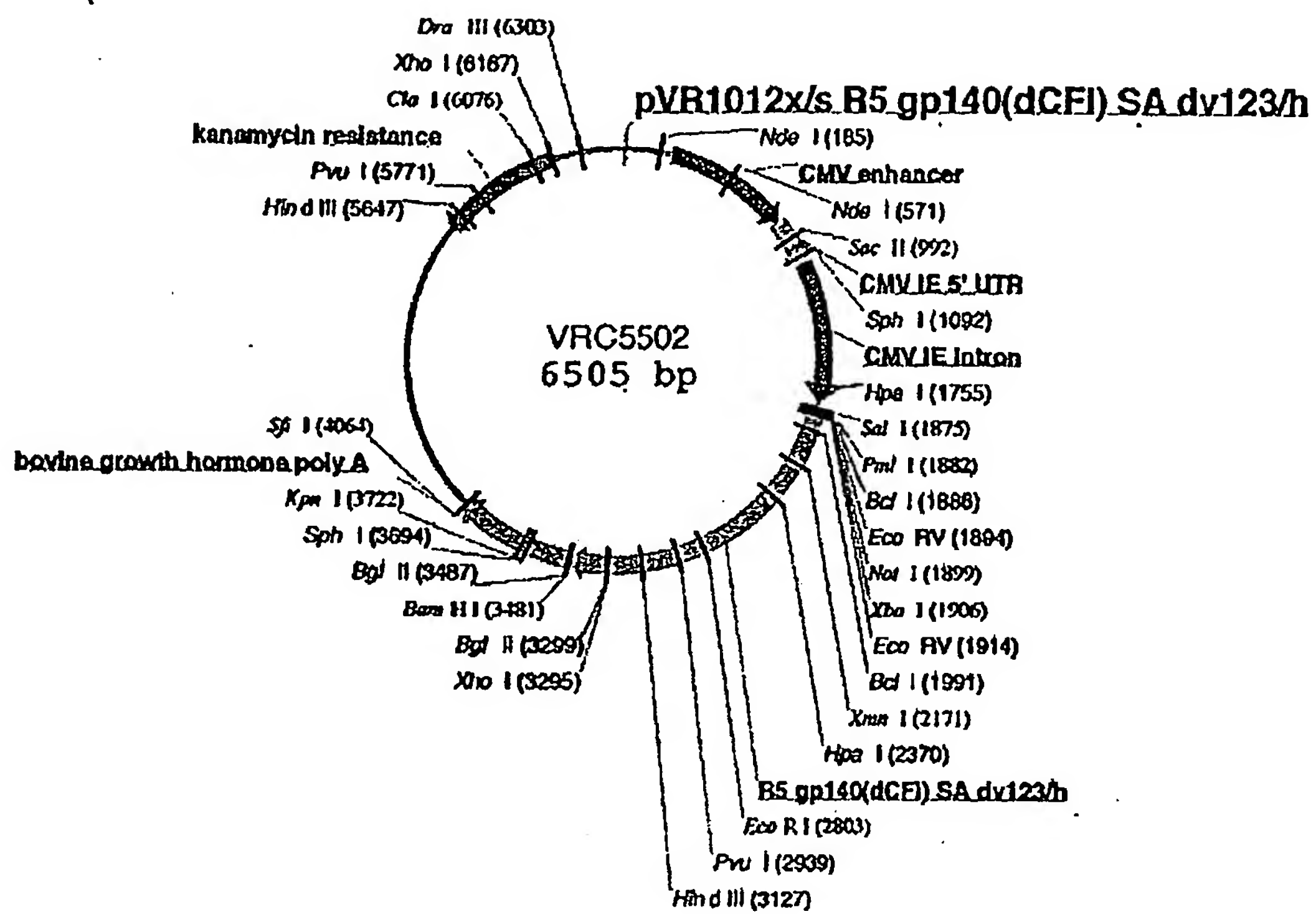


Figure 129

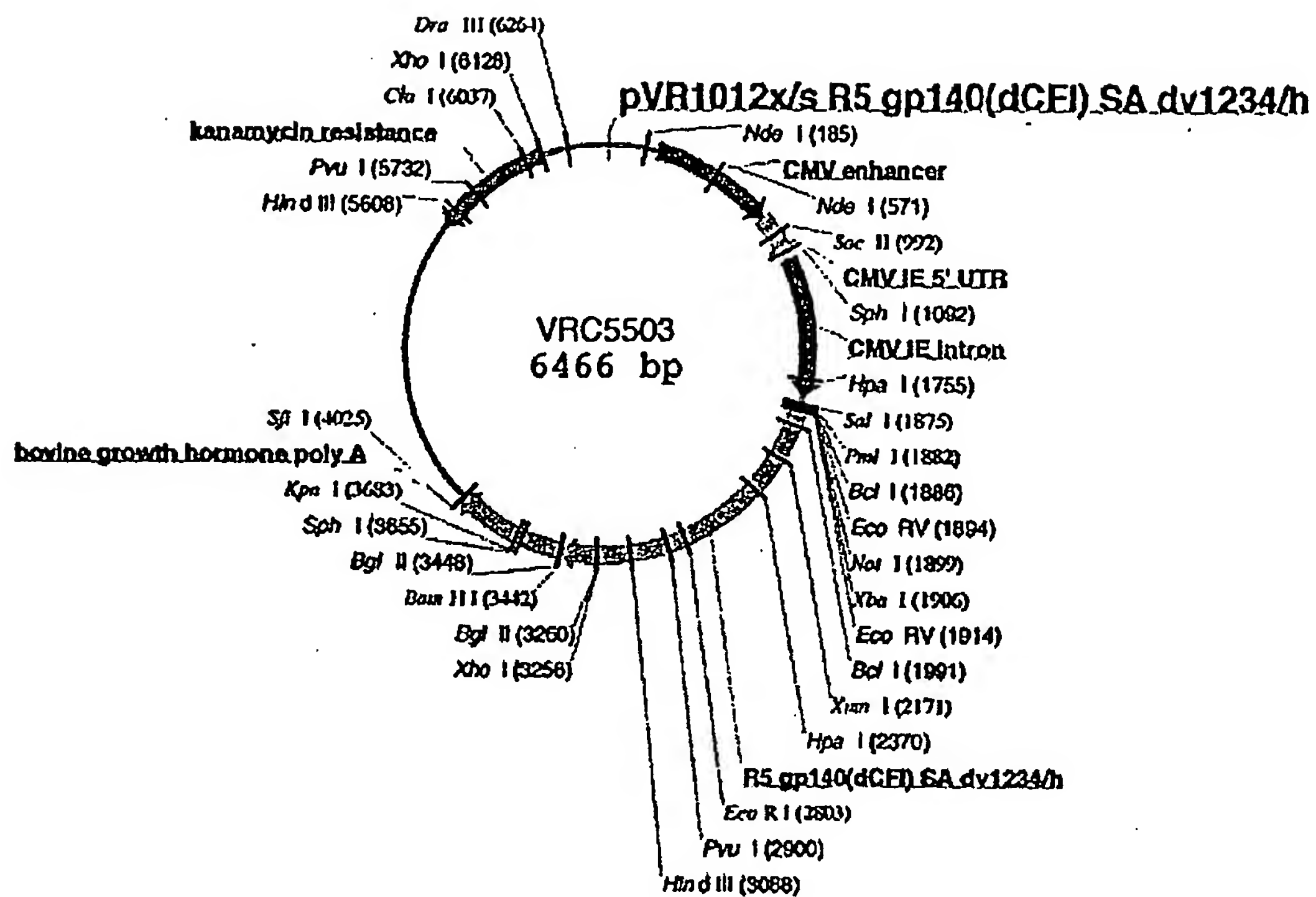


Figure 130

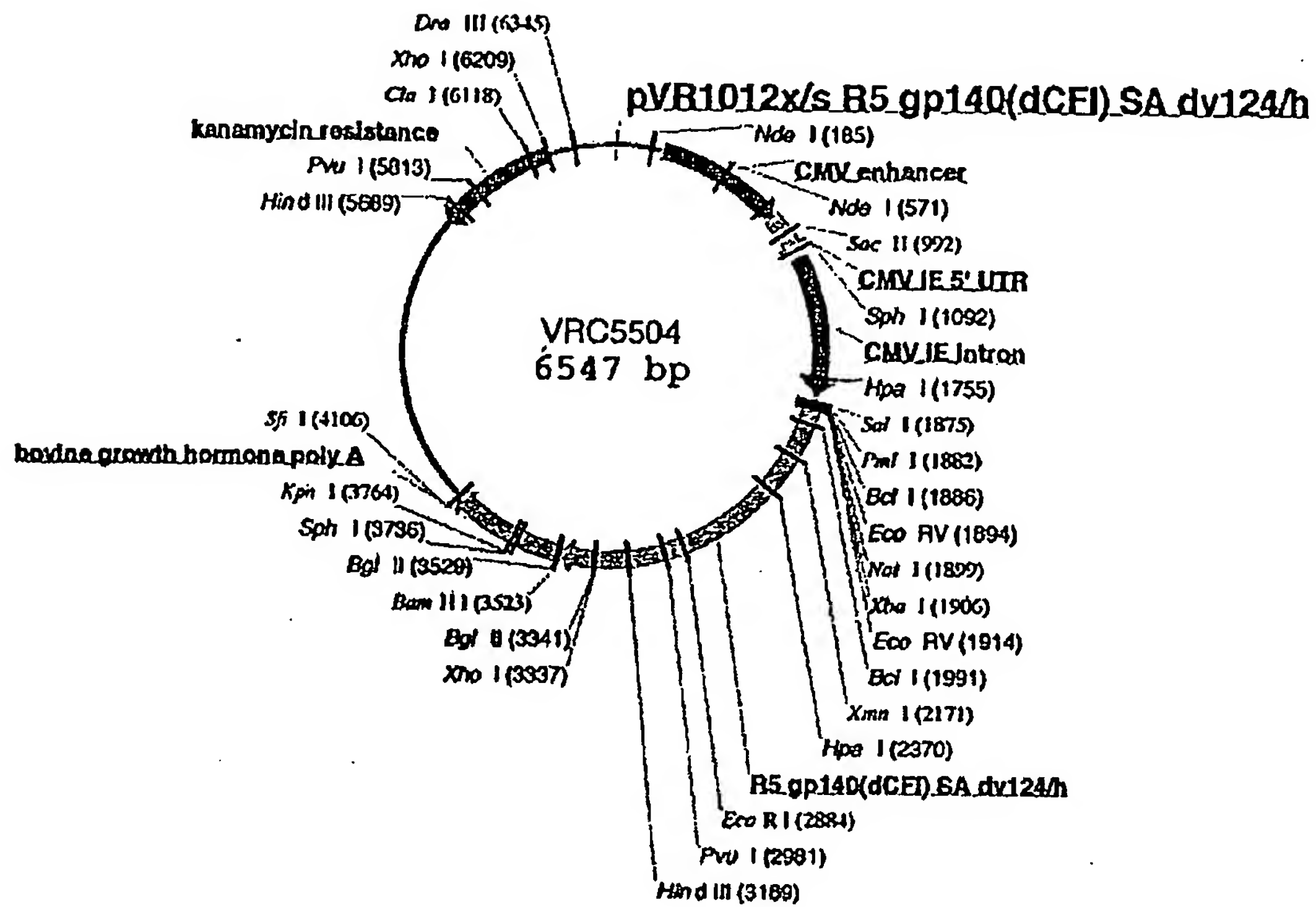


Figure 131

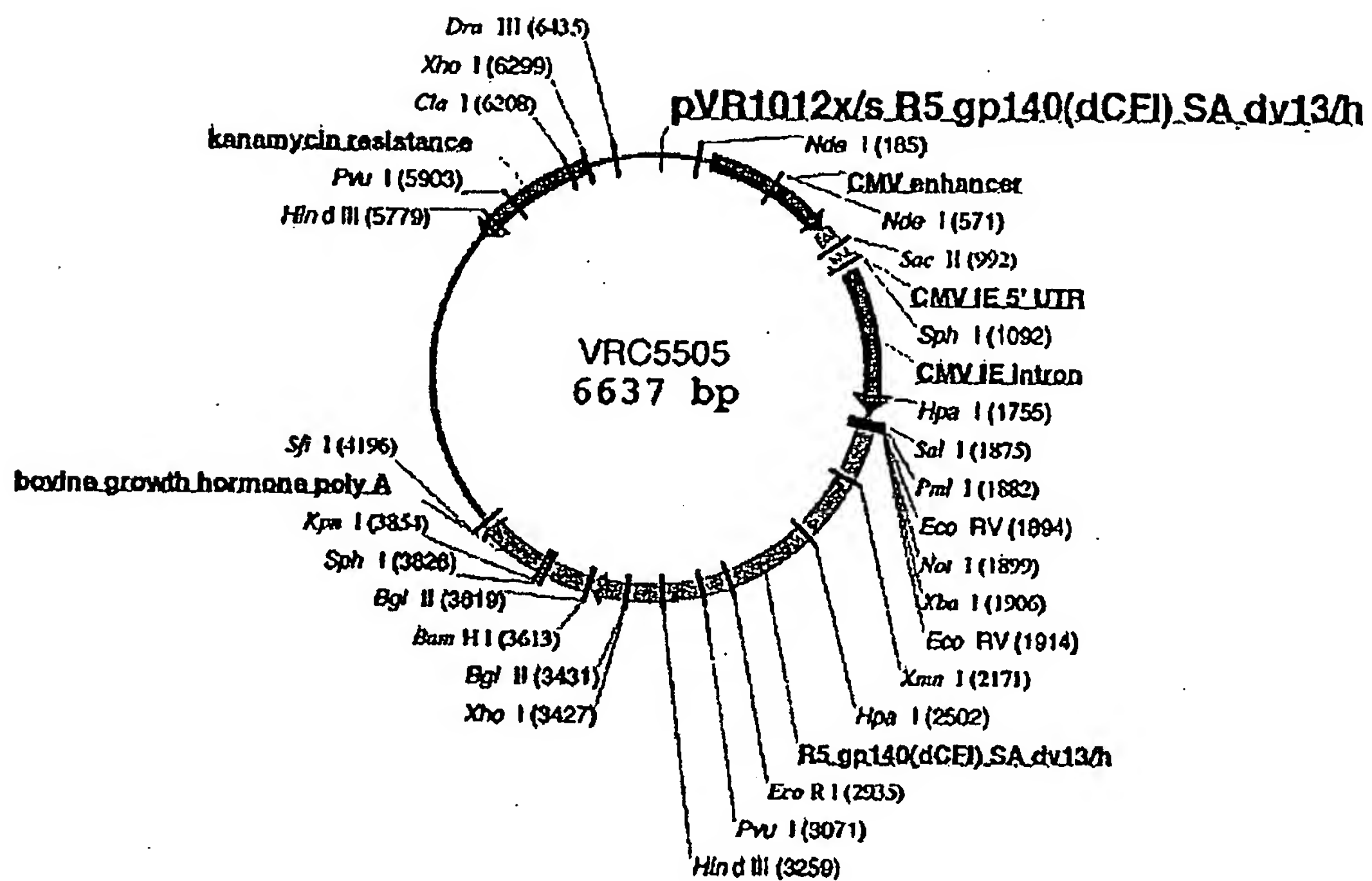


Figure 132



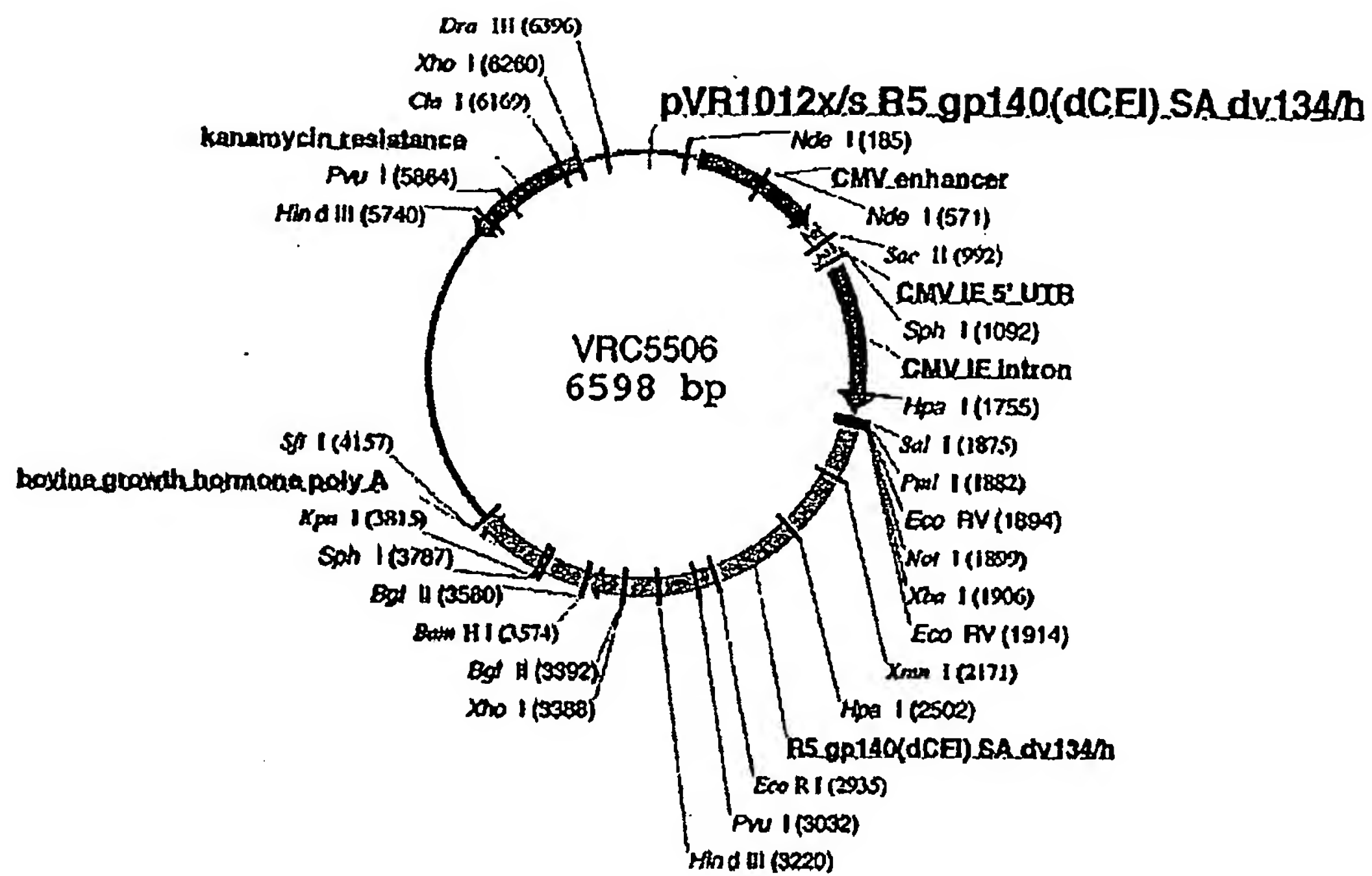


Figure 133

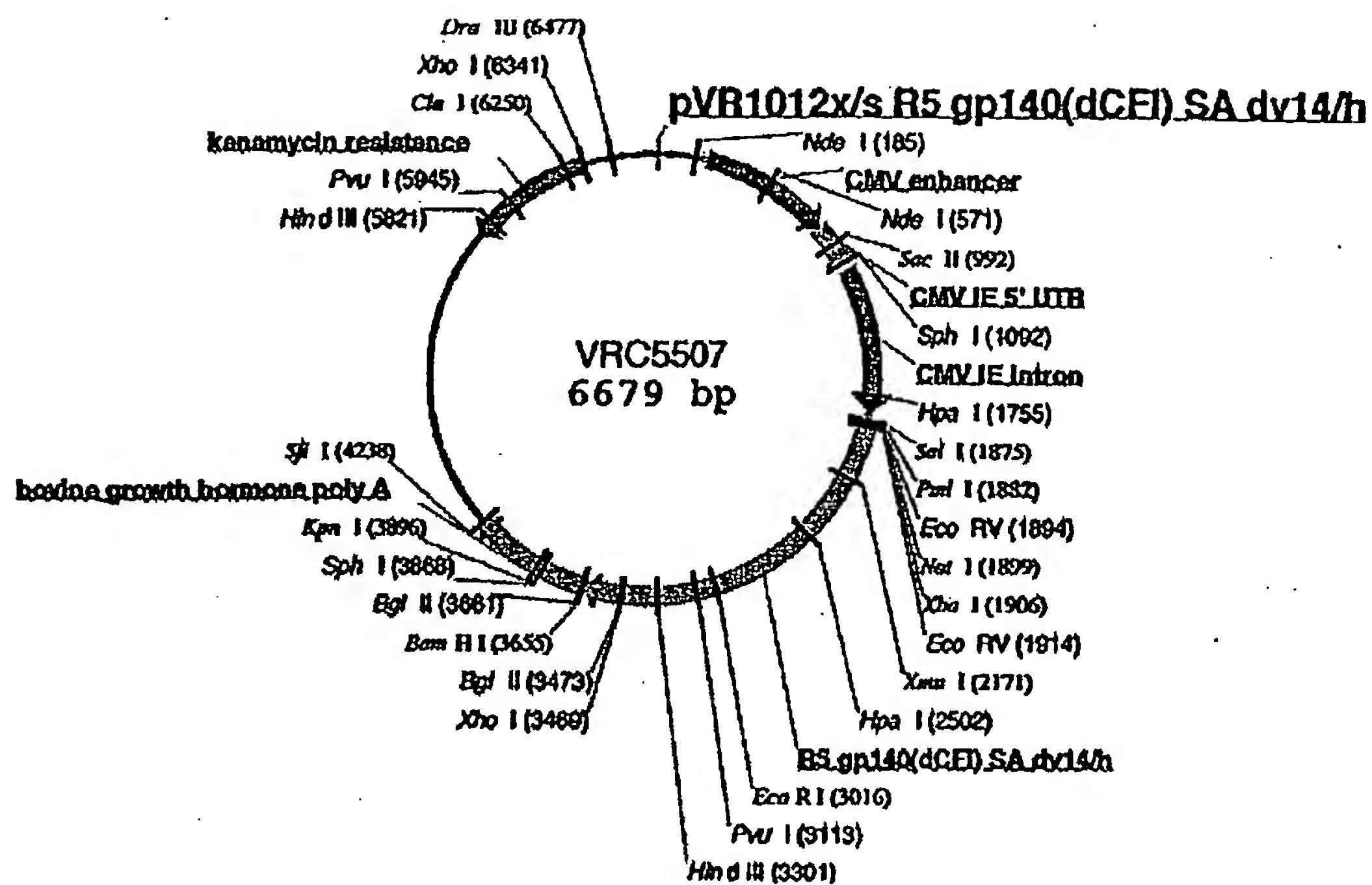


Figure 134

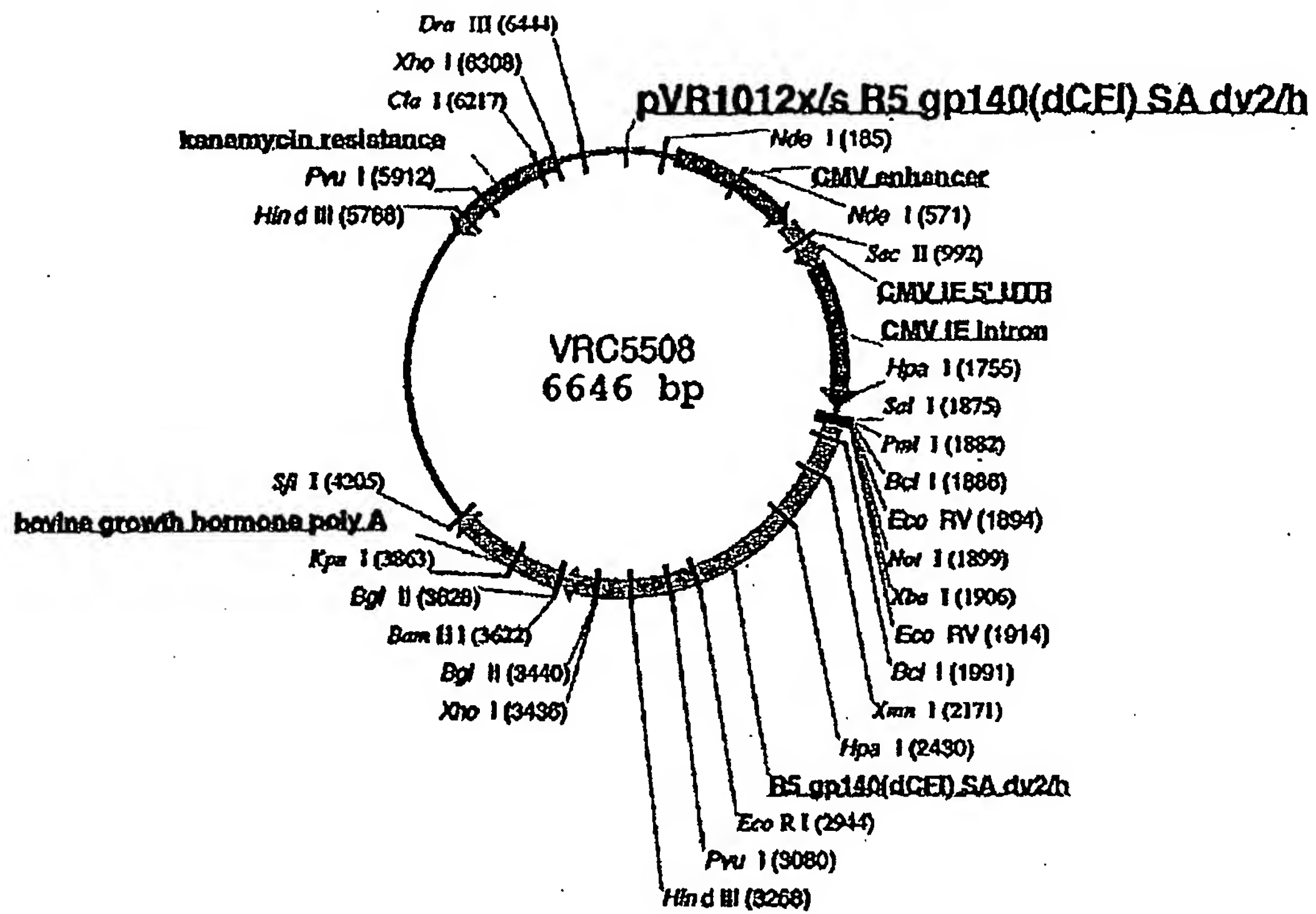


Figure 135

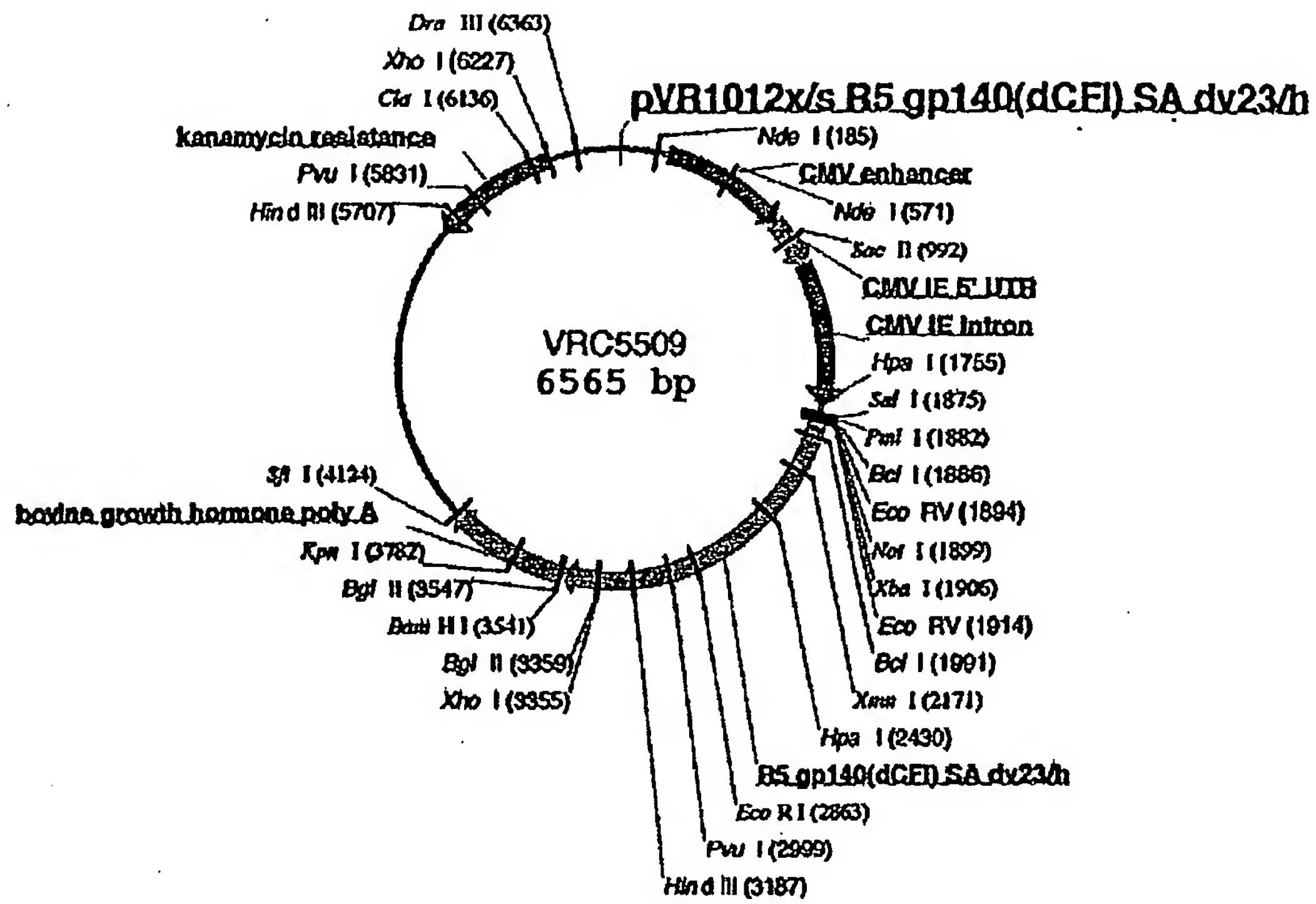


Figure 136

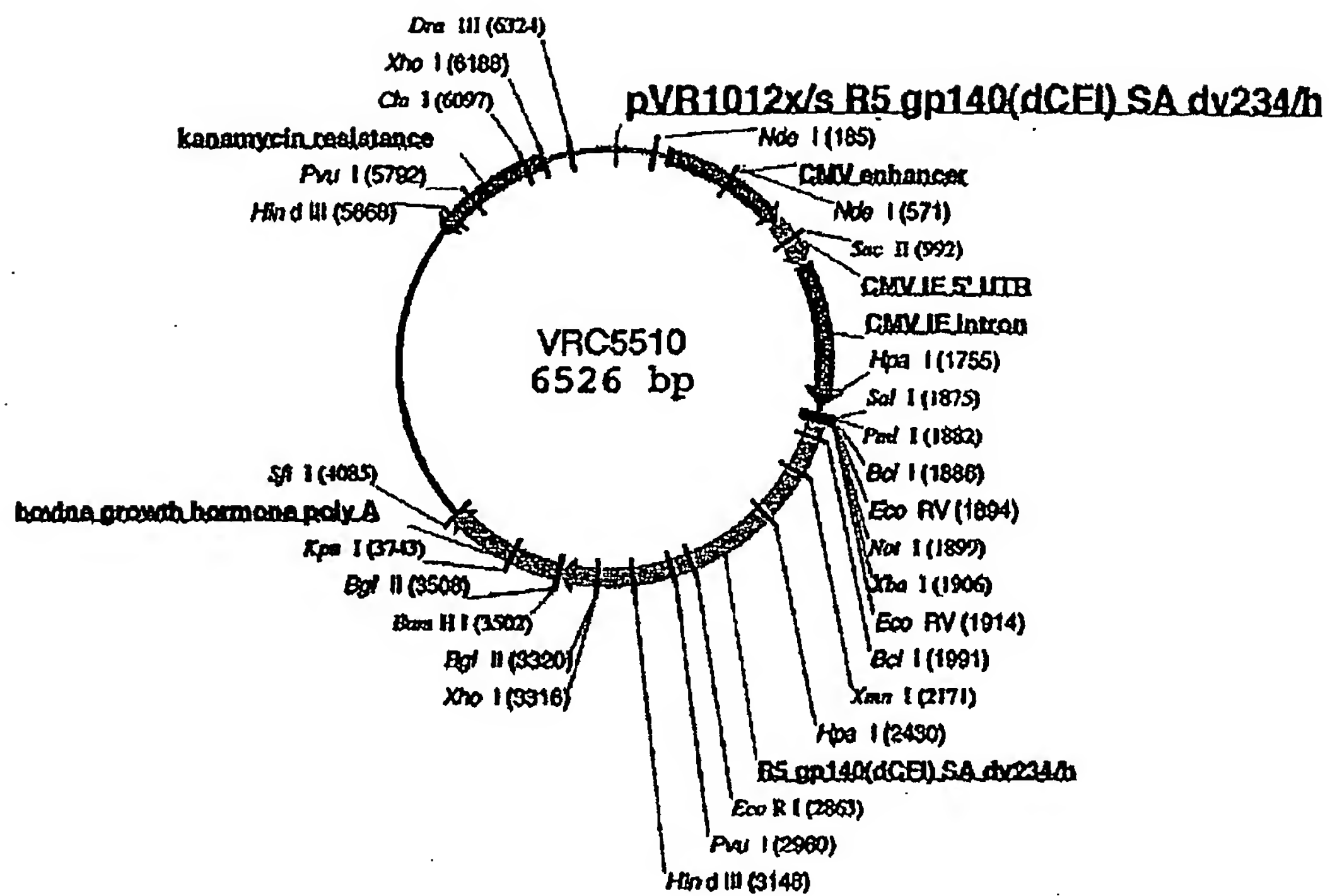


Figure 137

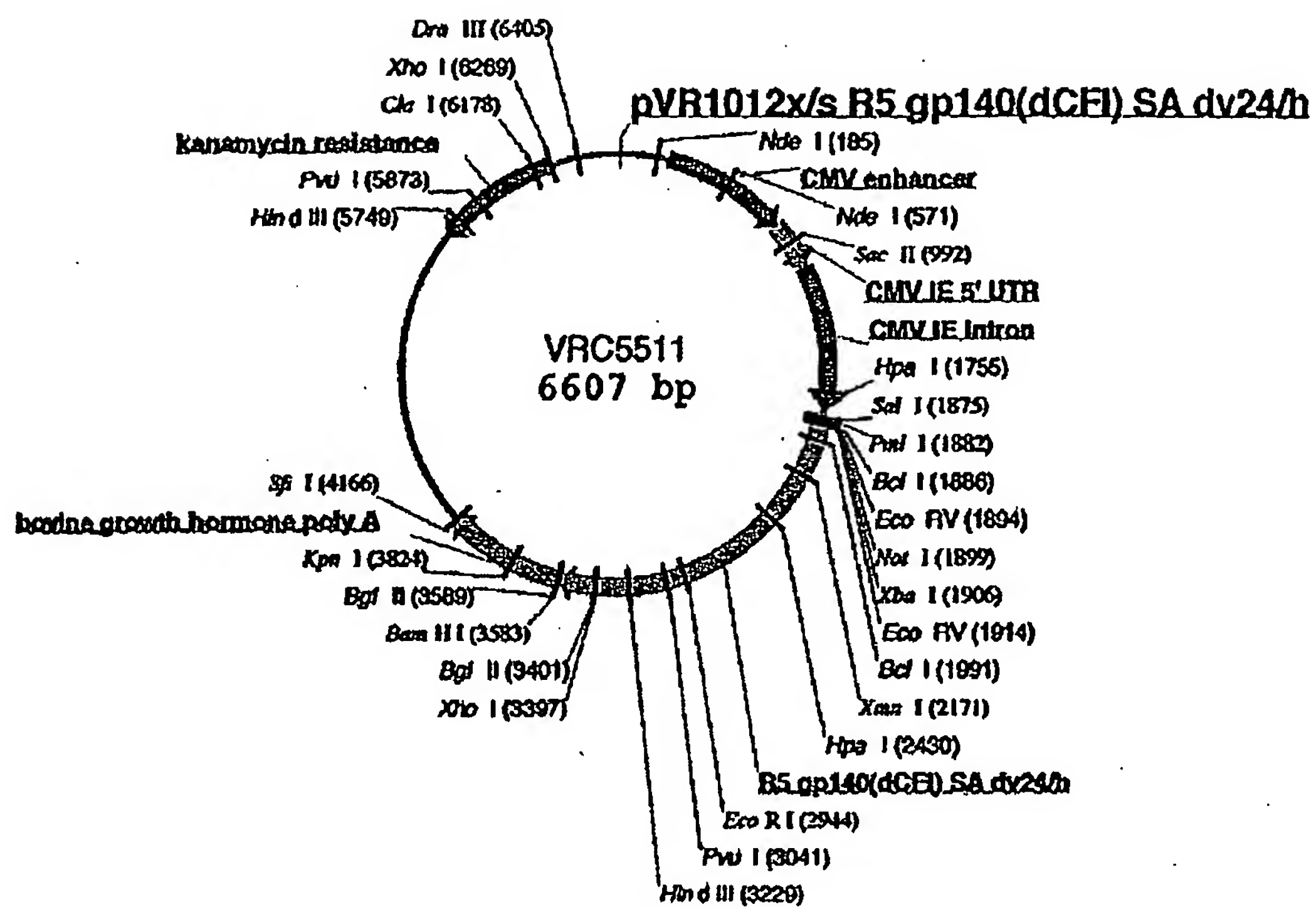


Figure 138

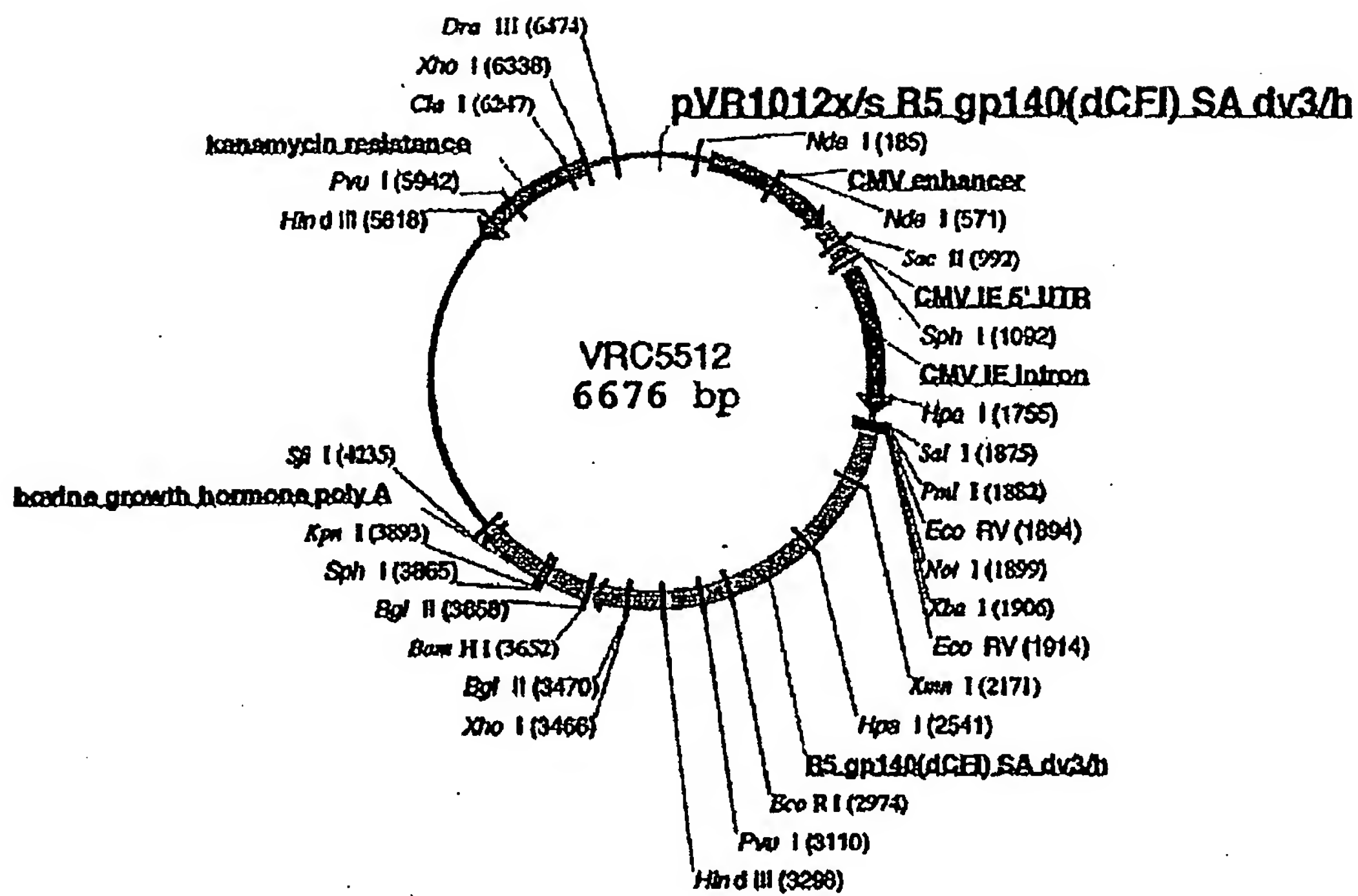


Figure 139



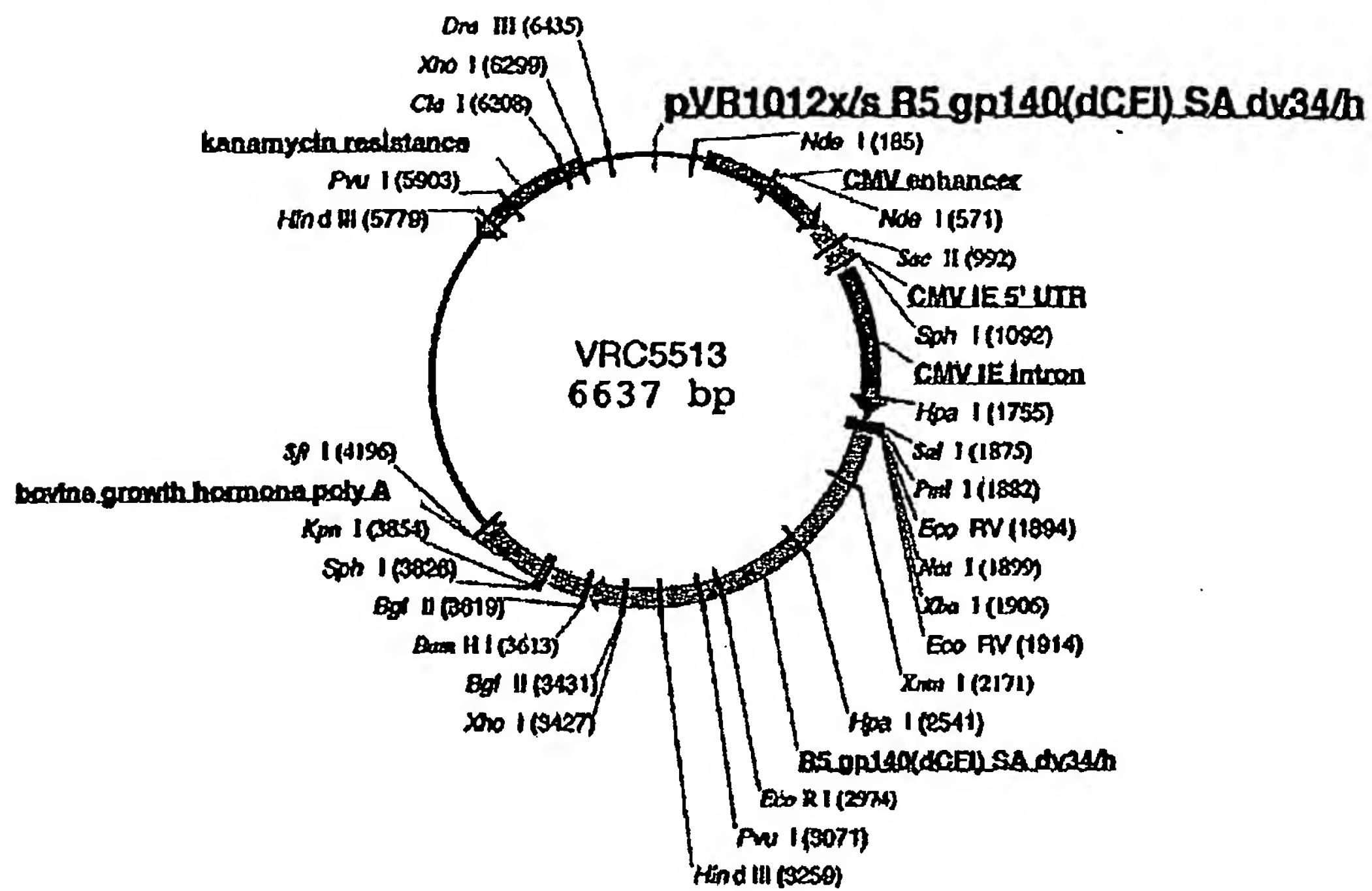


Figure 140

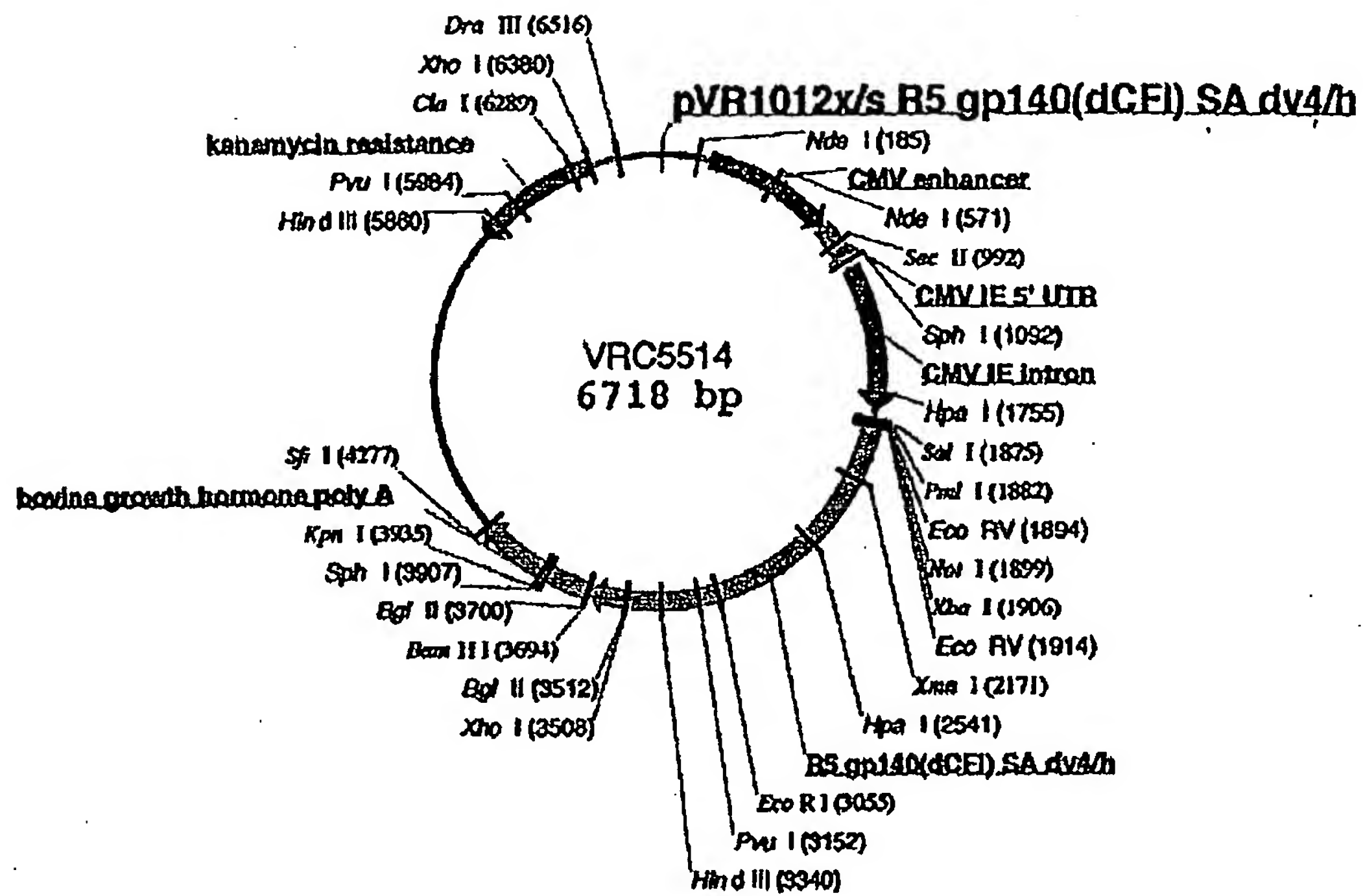


Figure 141

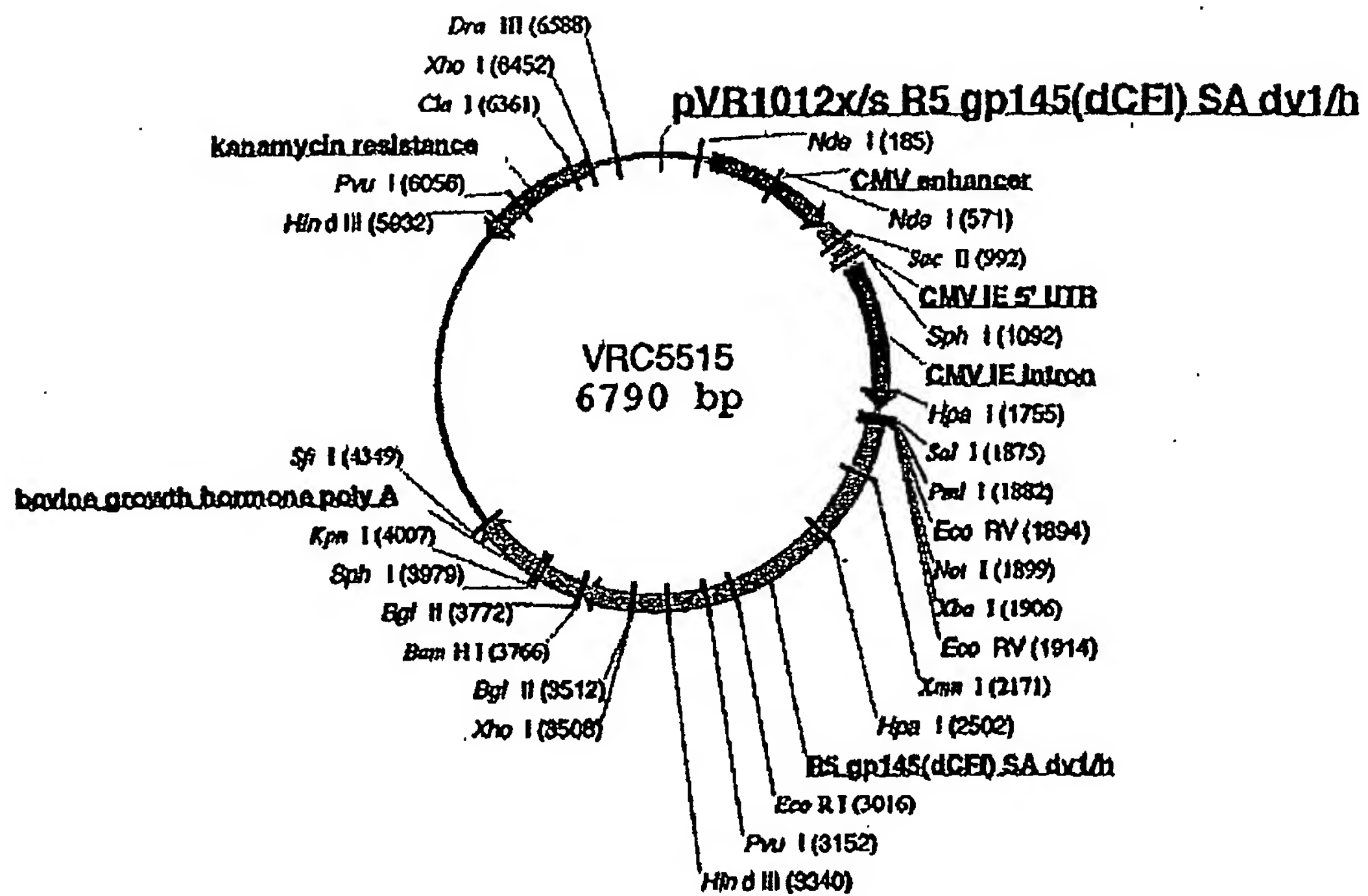


Figure 142

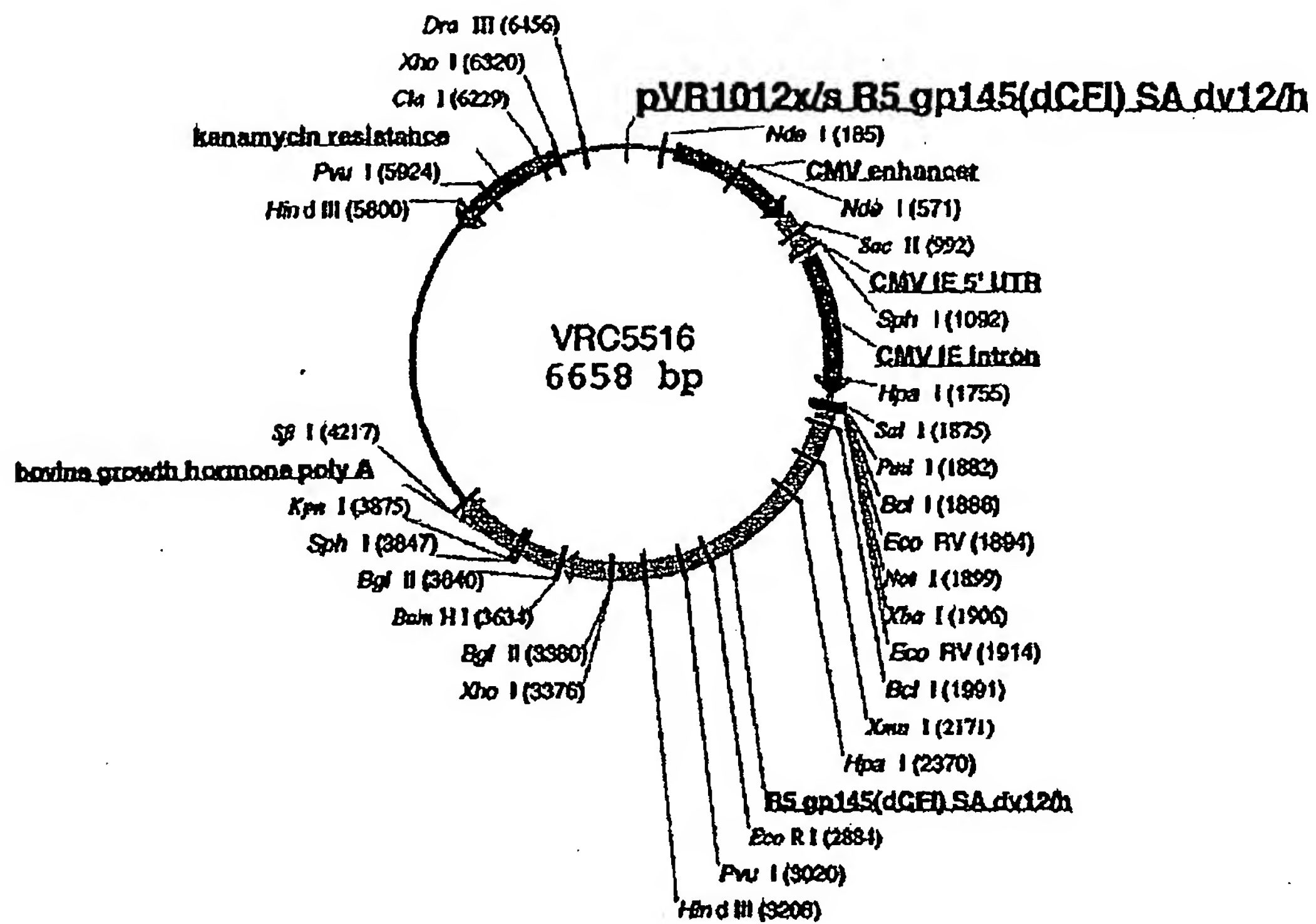


Figure 143

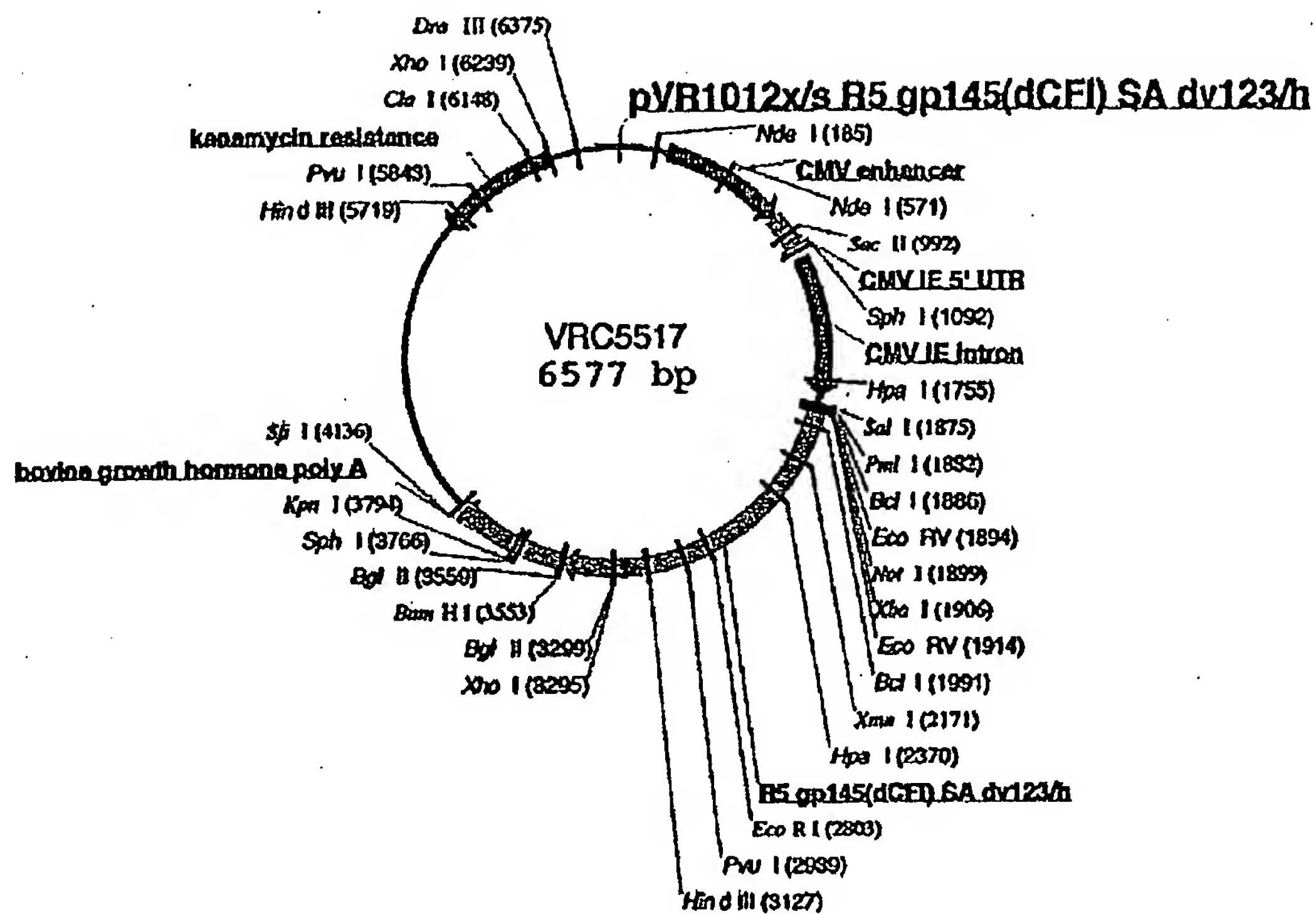


Figure 144

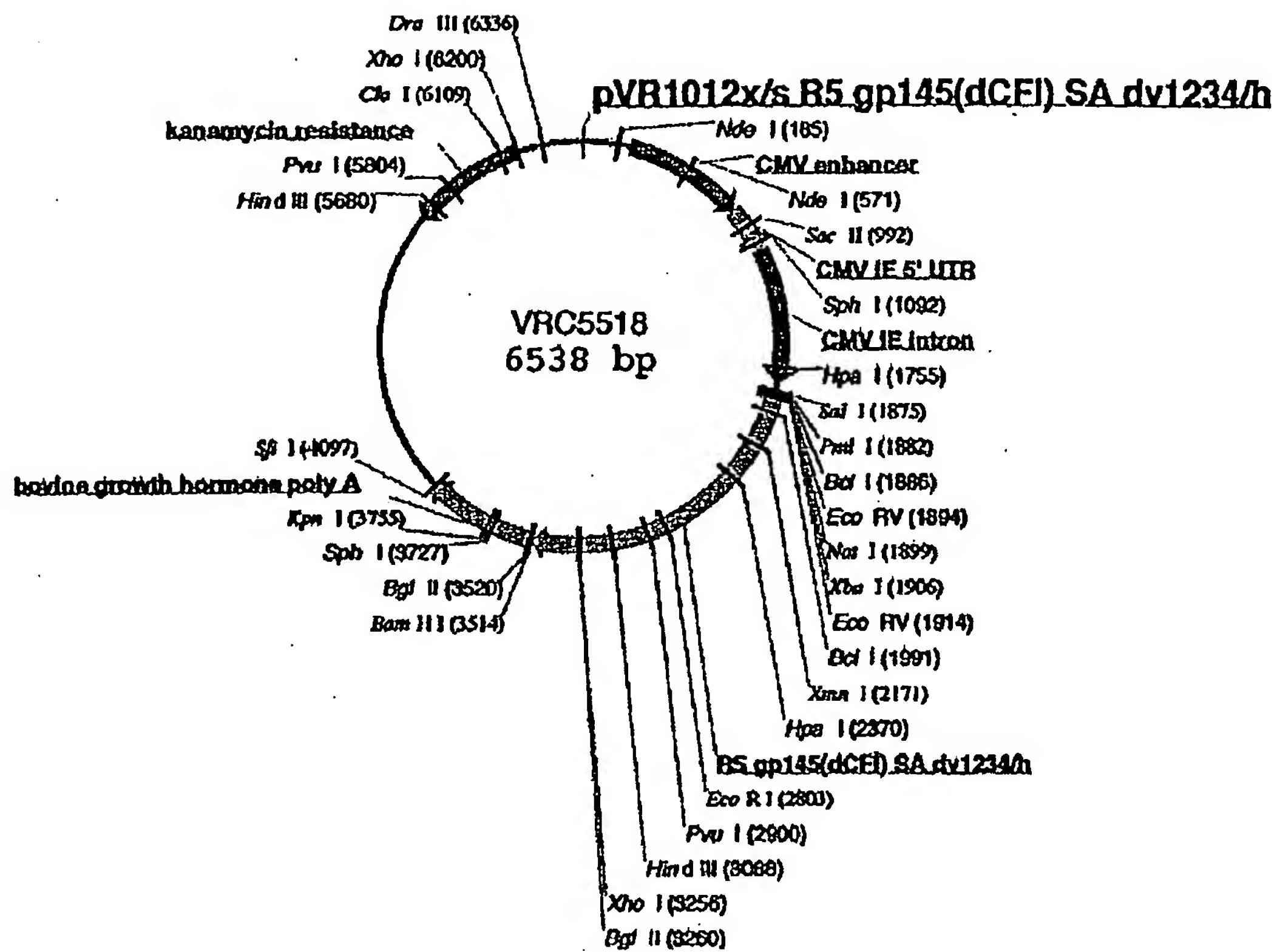


Figure 145

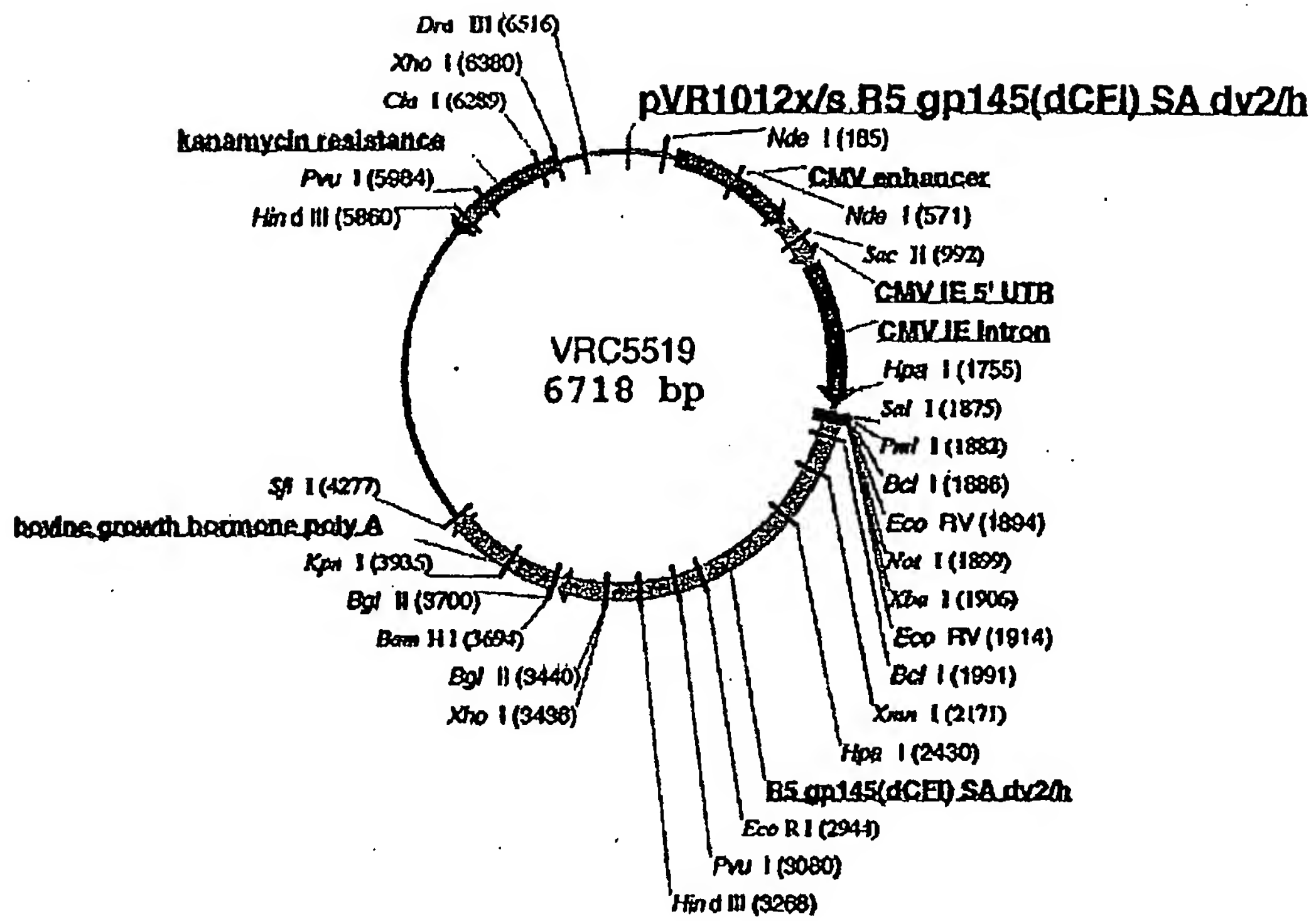


Figure 146



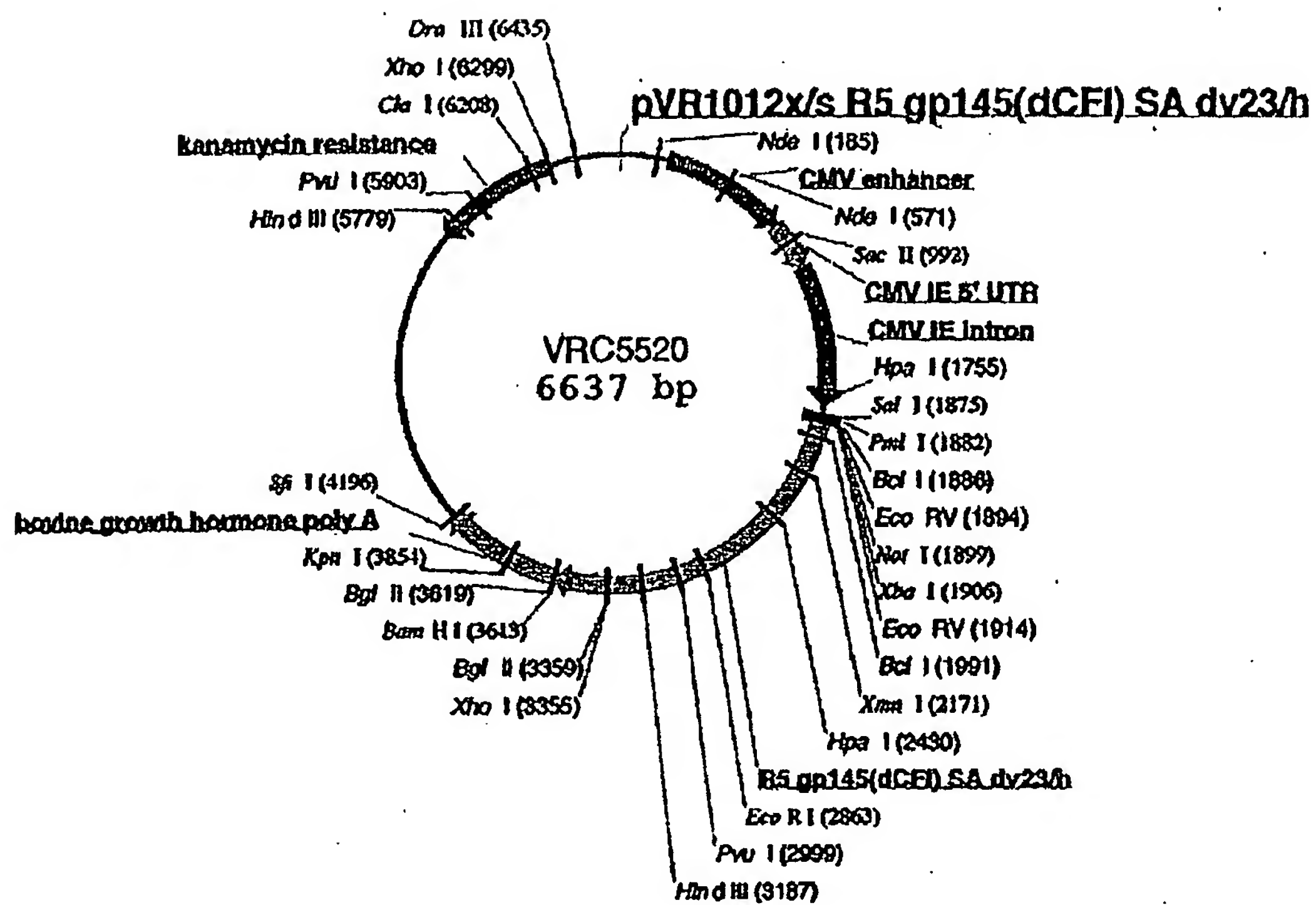


Figure 147

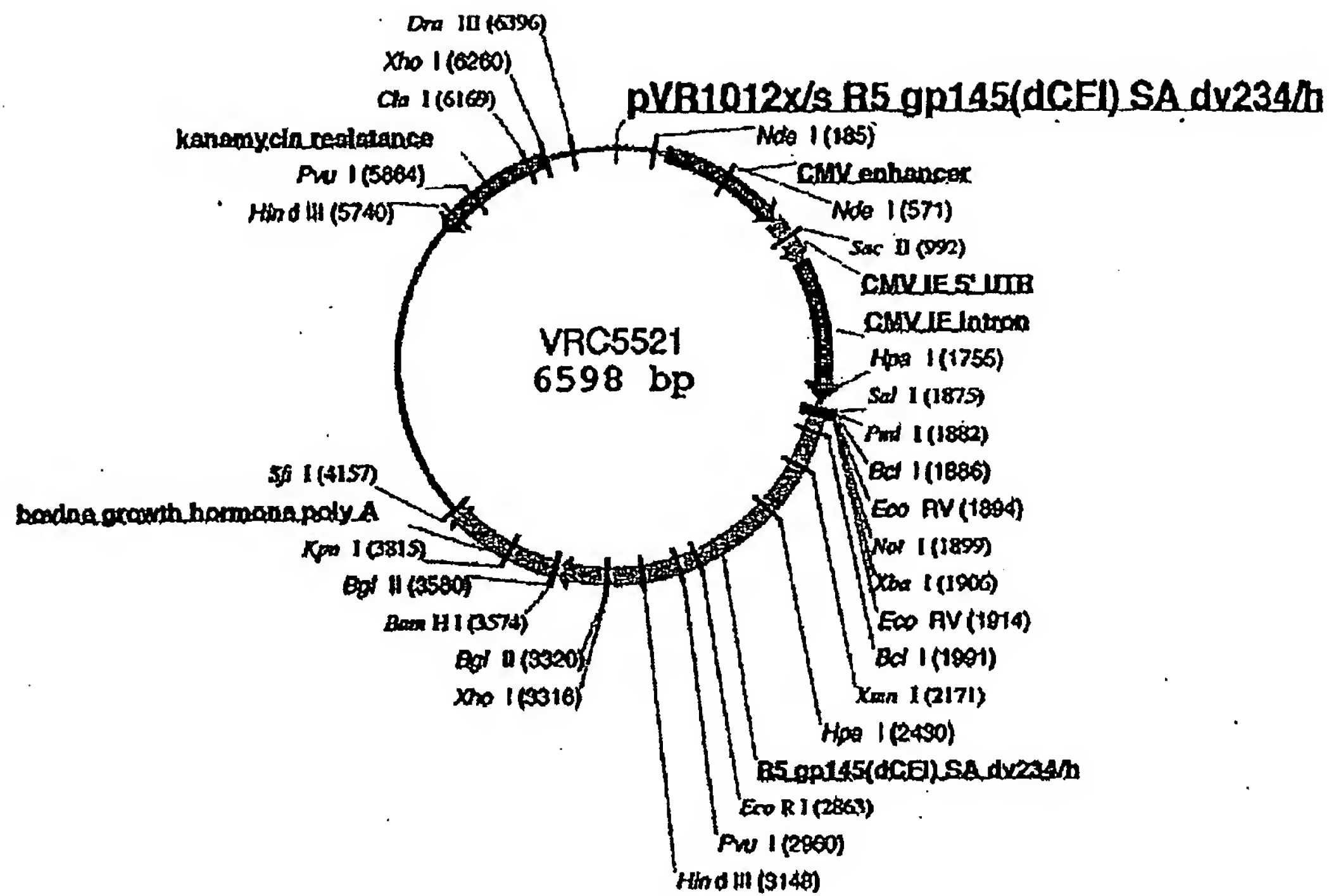


Figure 148

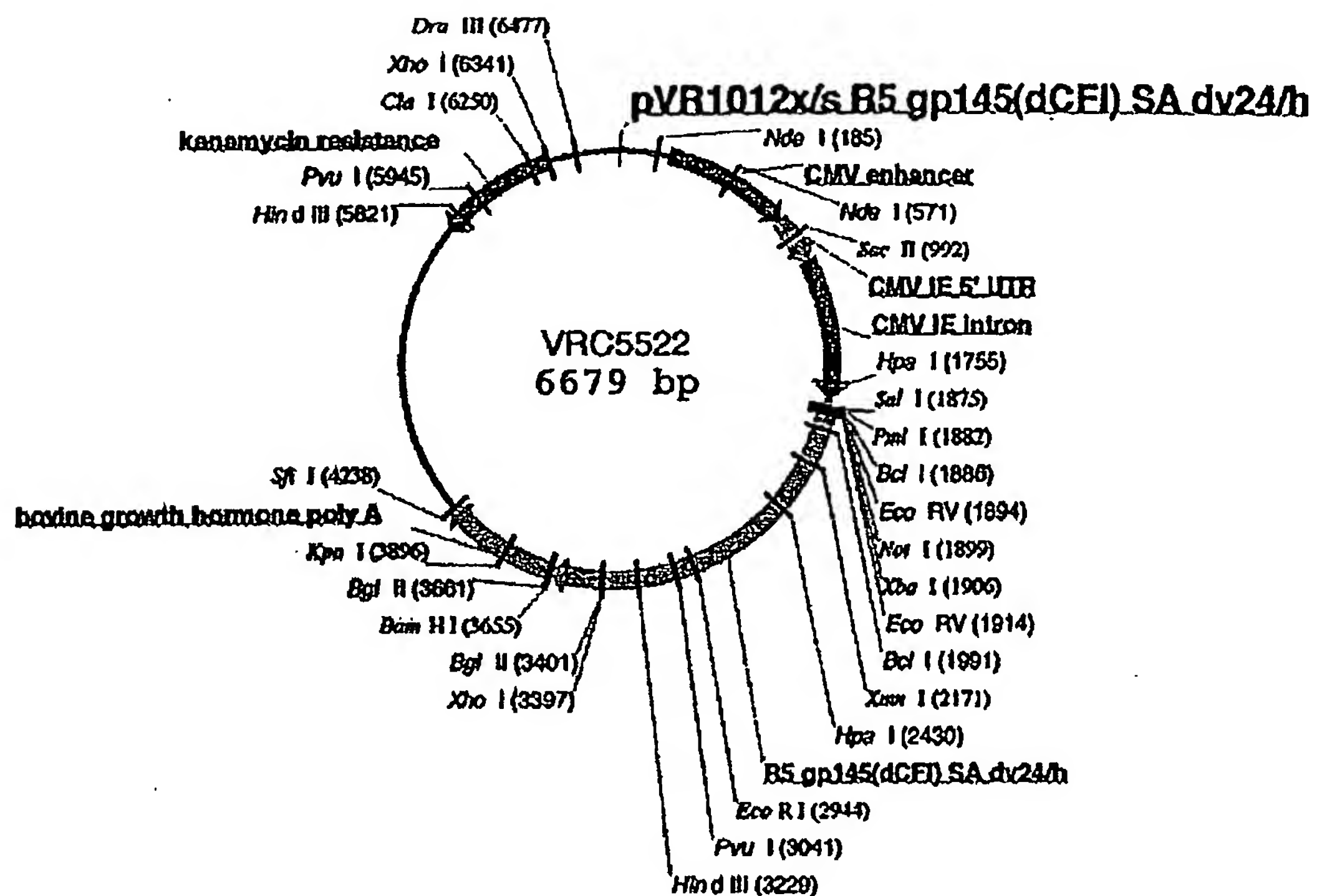


Figure 149

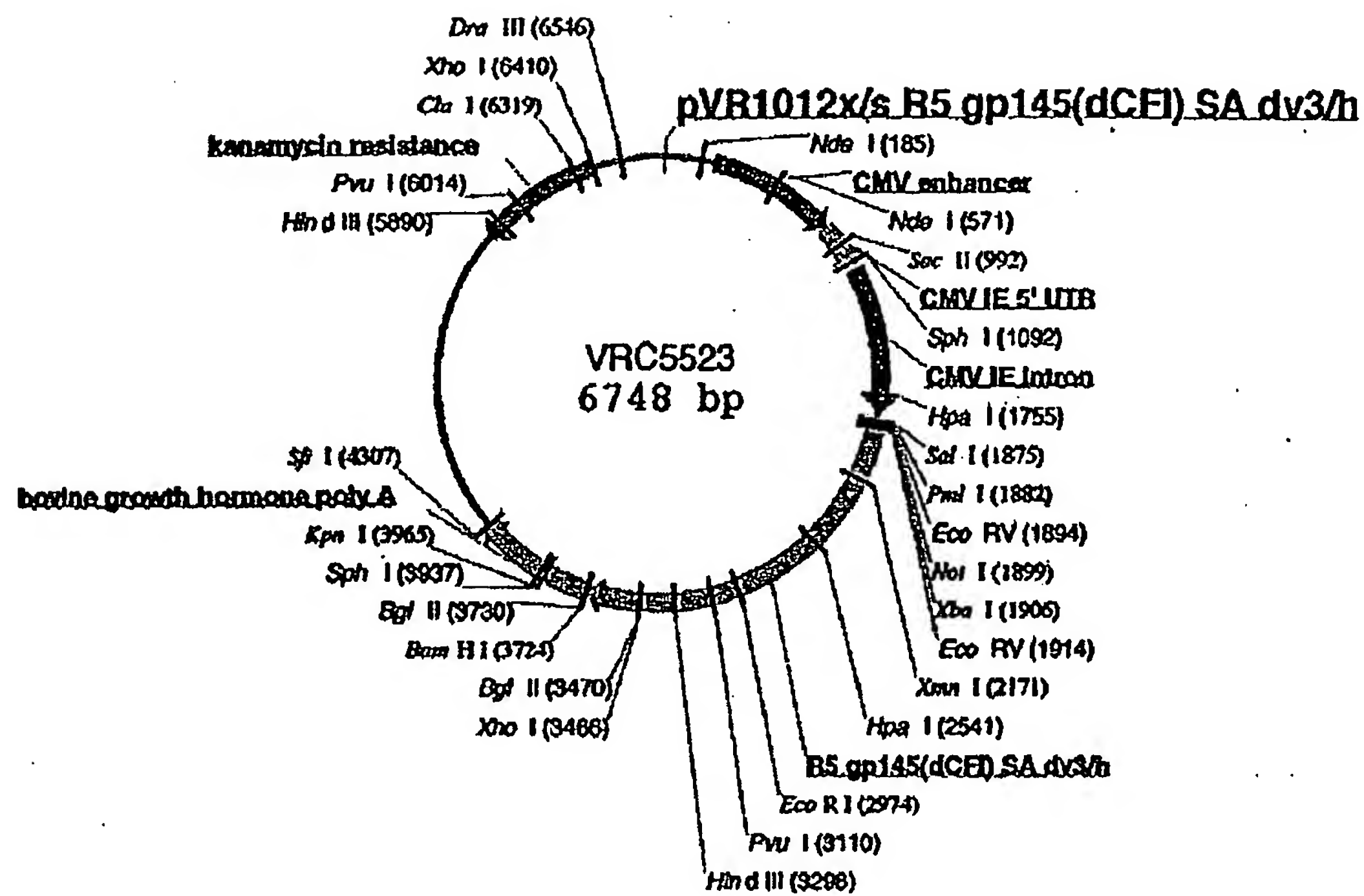


Figure 150

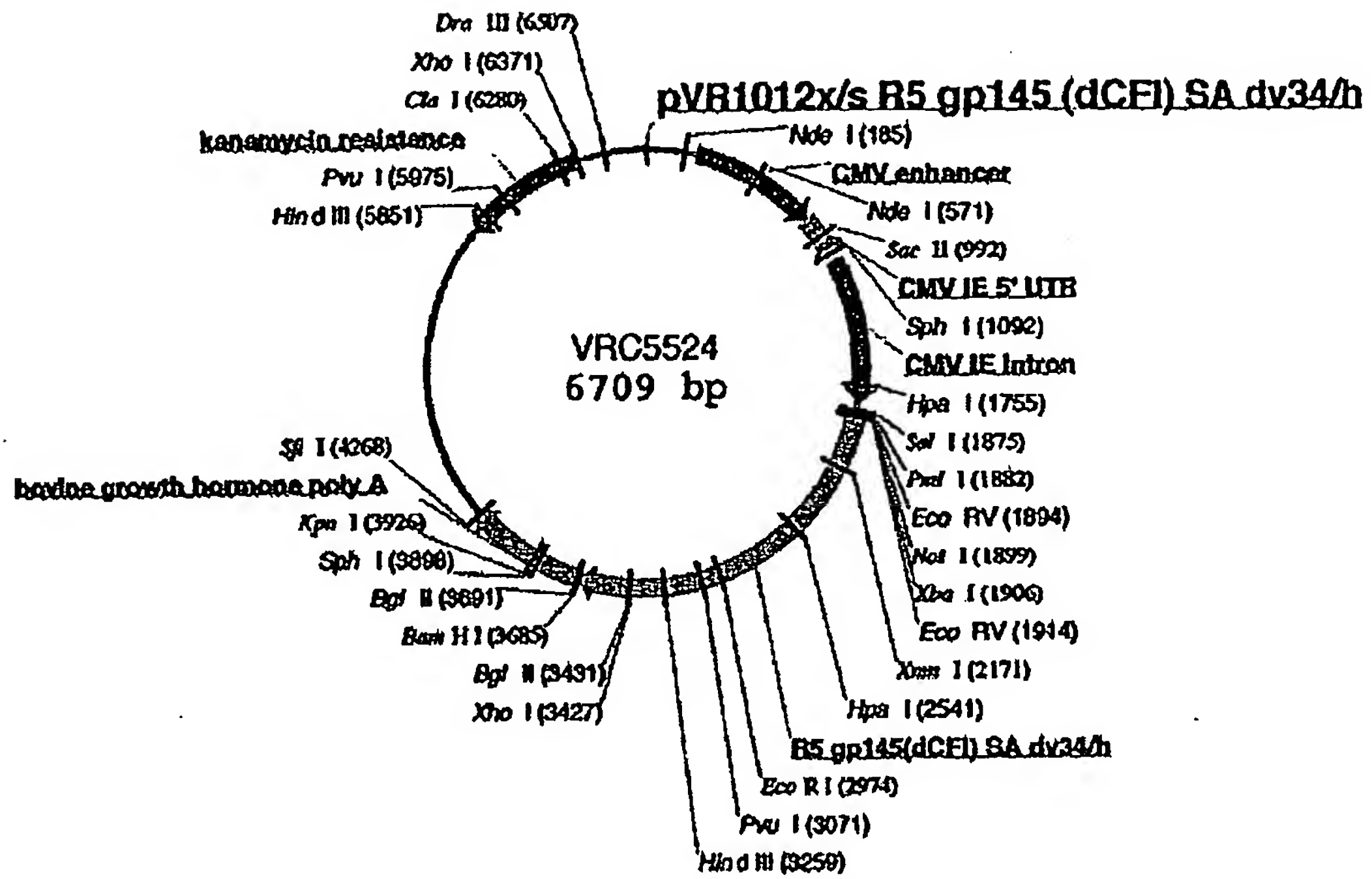


Figure 151

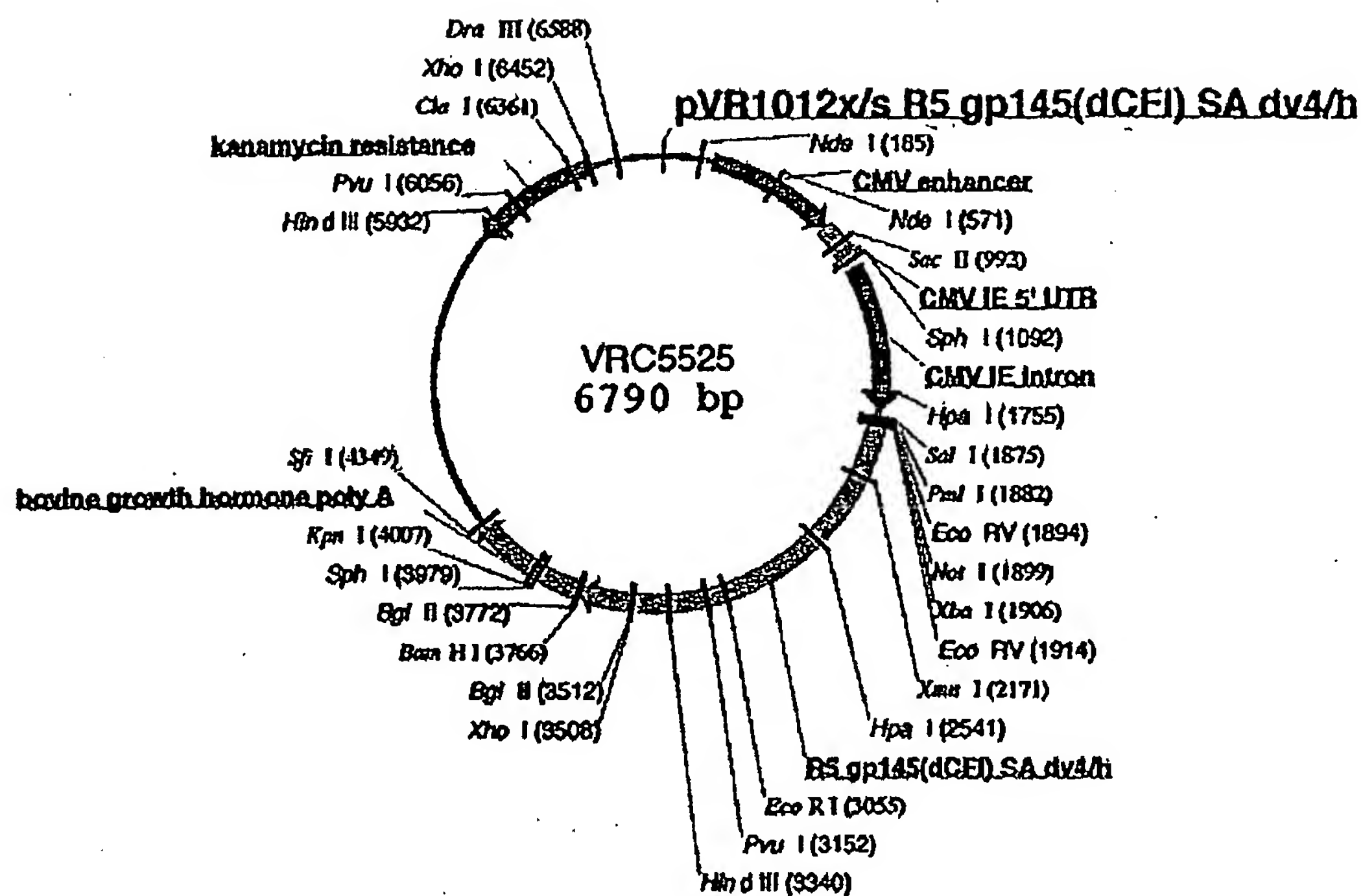


Figure 152

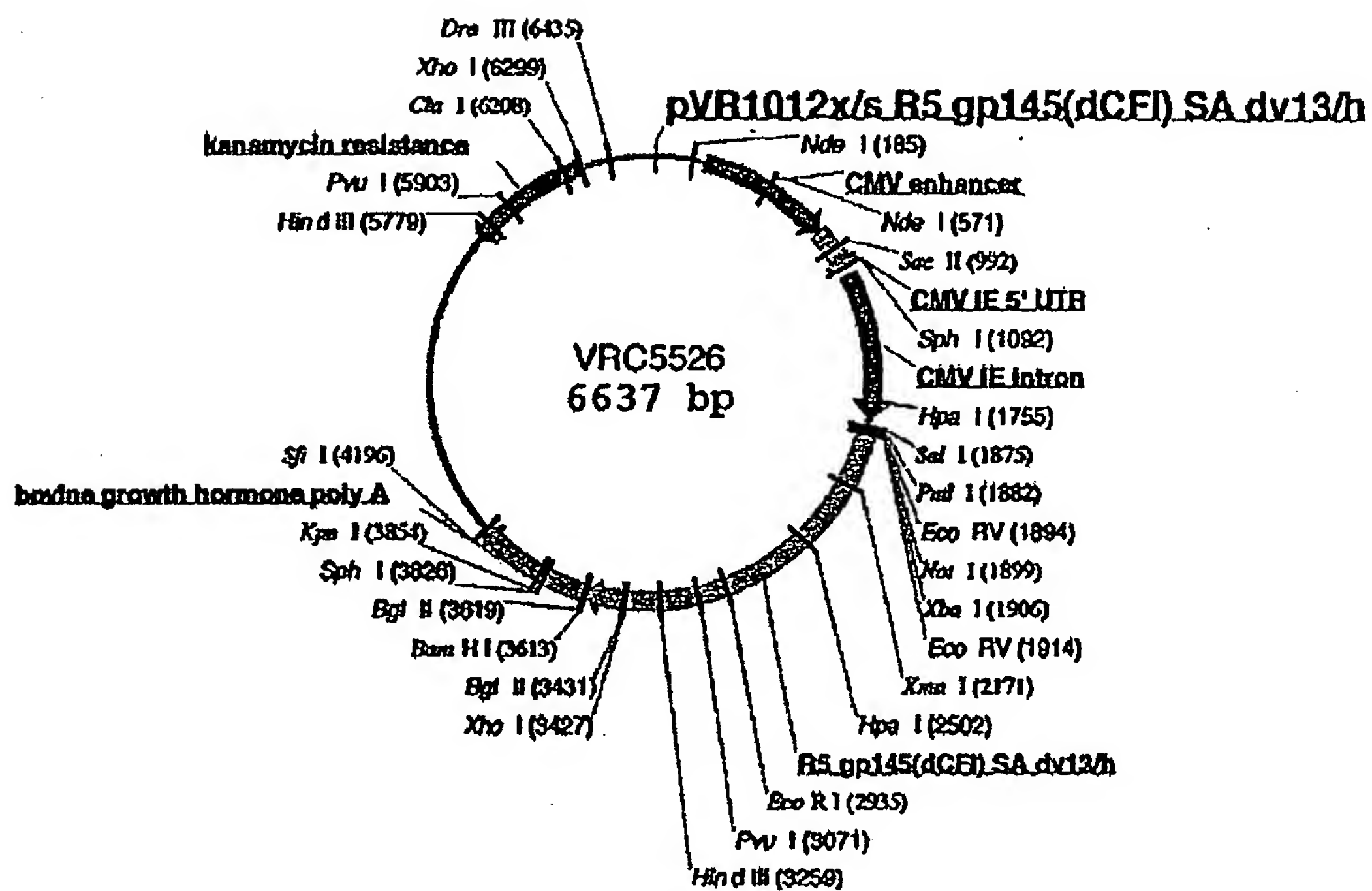


Figure 153



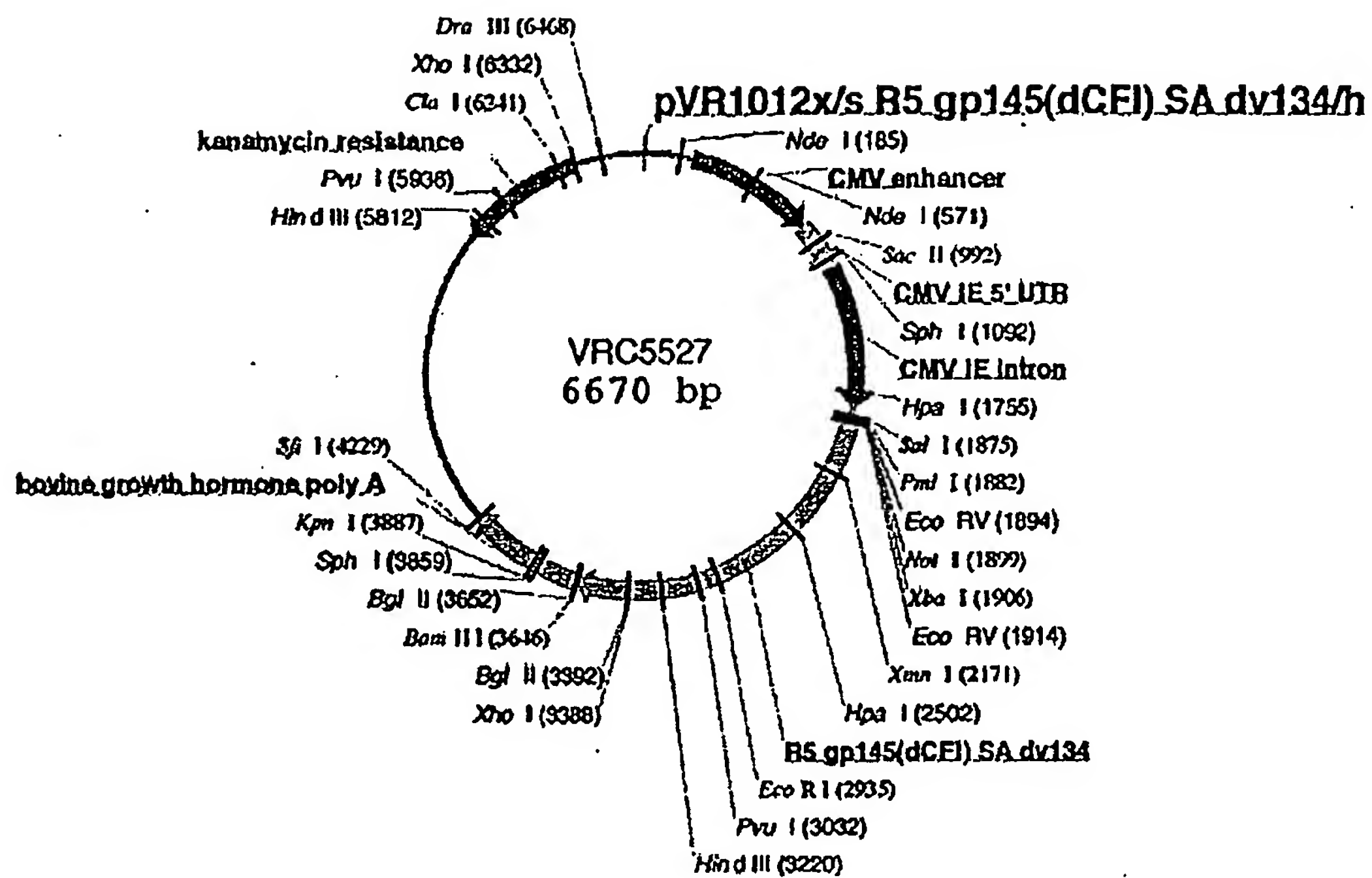


Figure 154

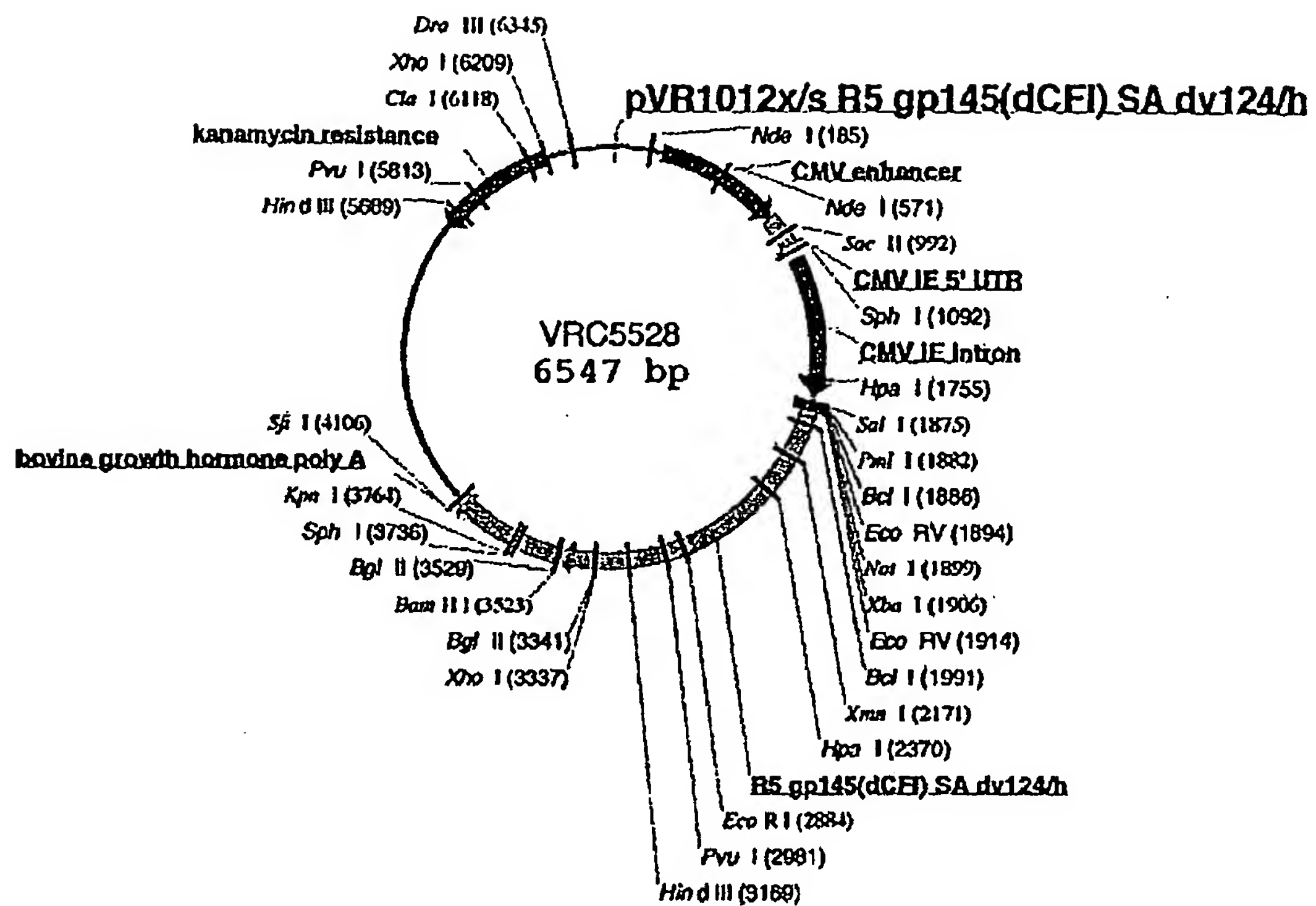


Figure 155

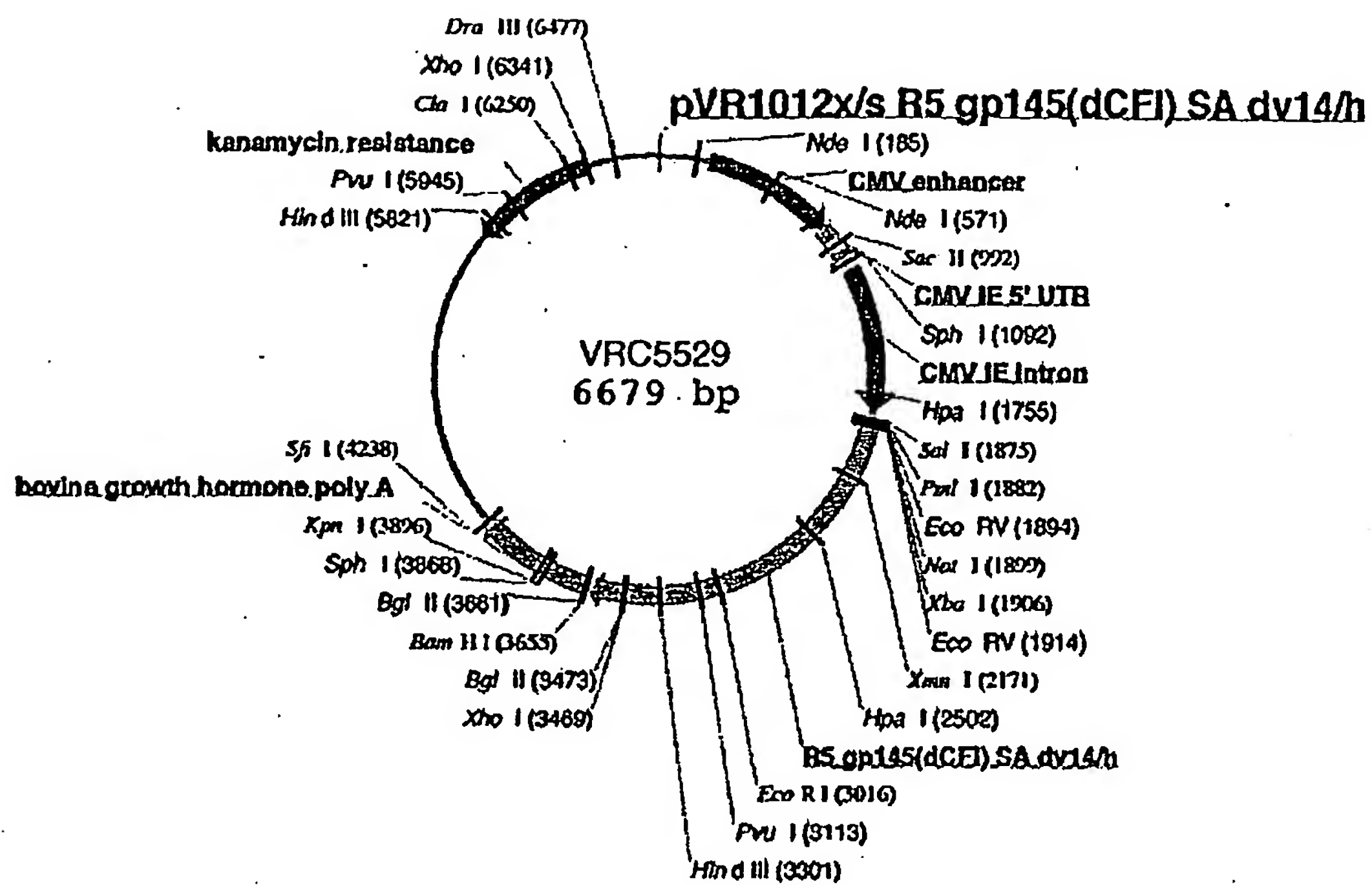


Figure 156

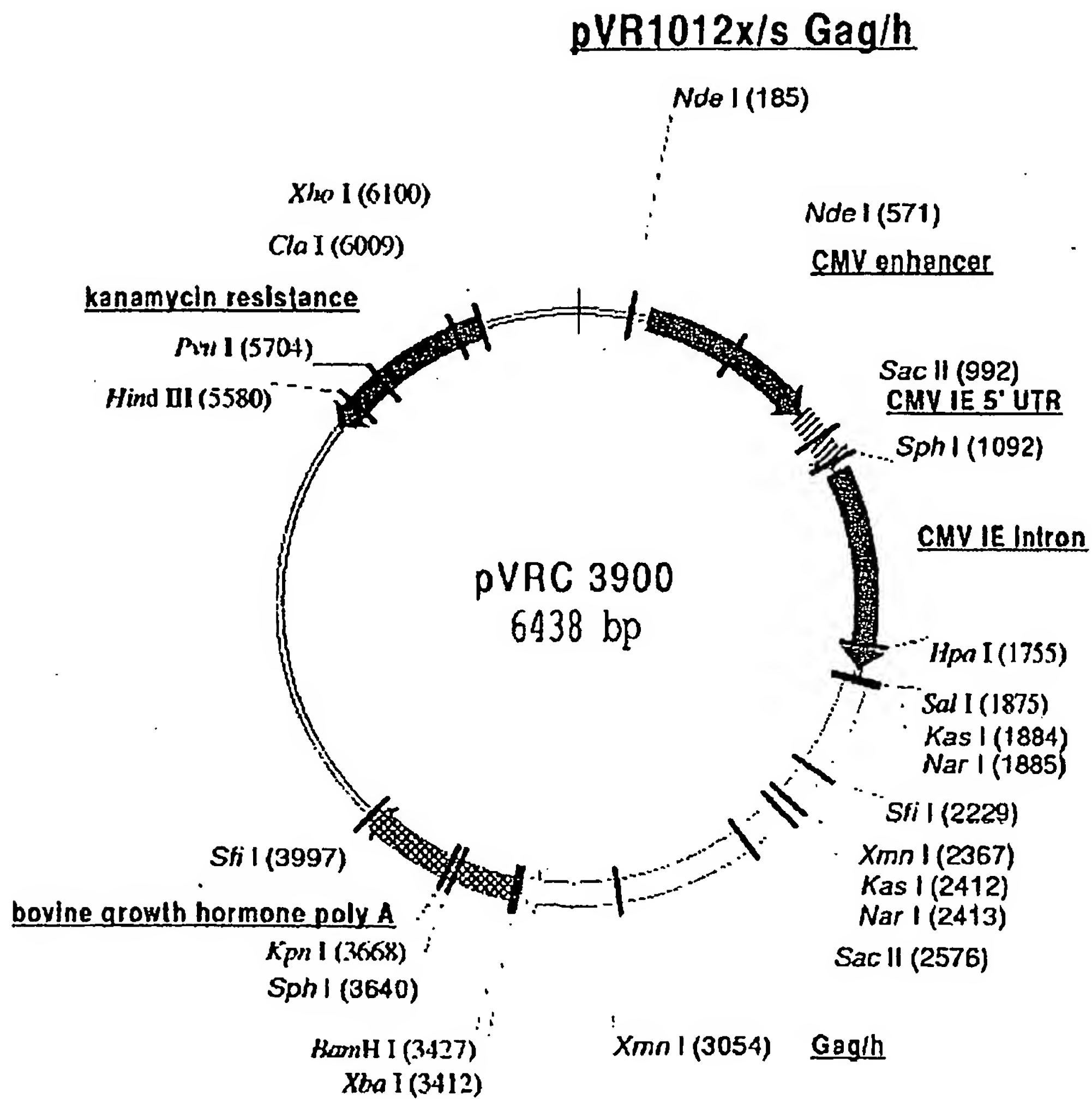


Figure 157

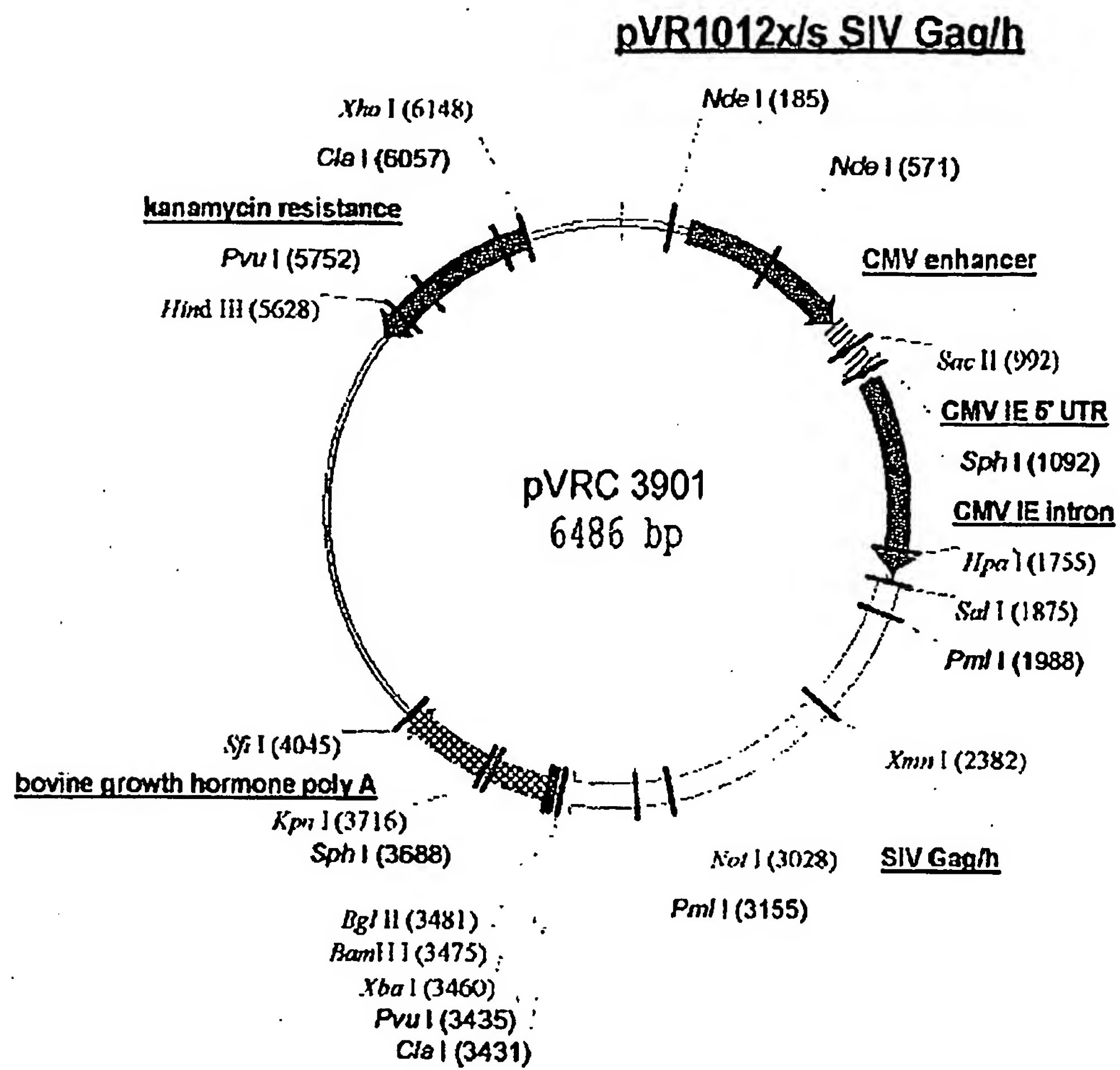


Figure 158

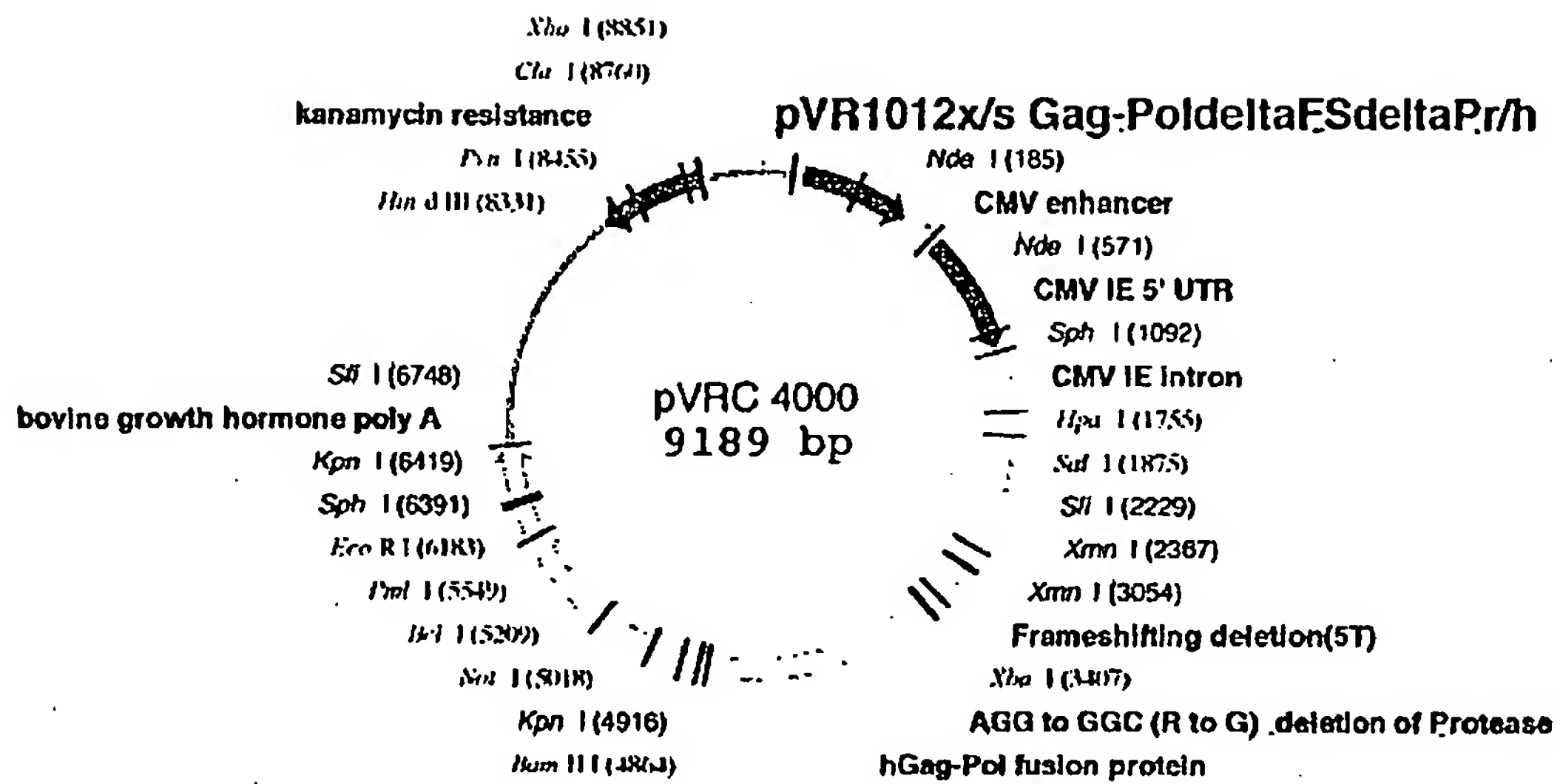


Figure 159

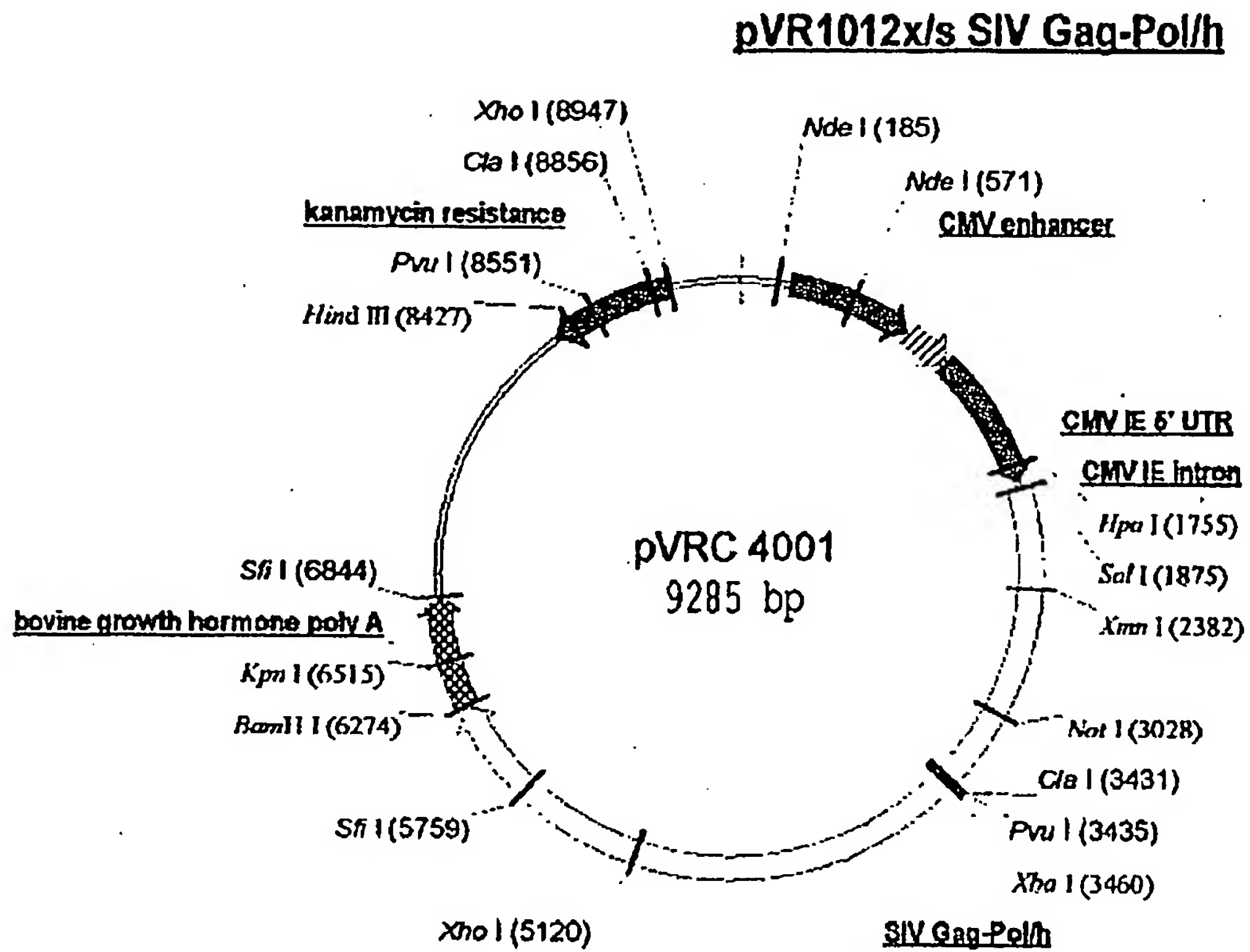


Figure 160



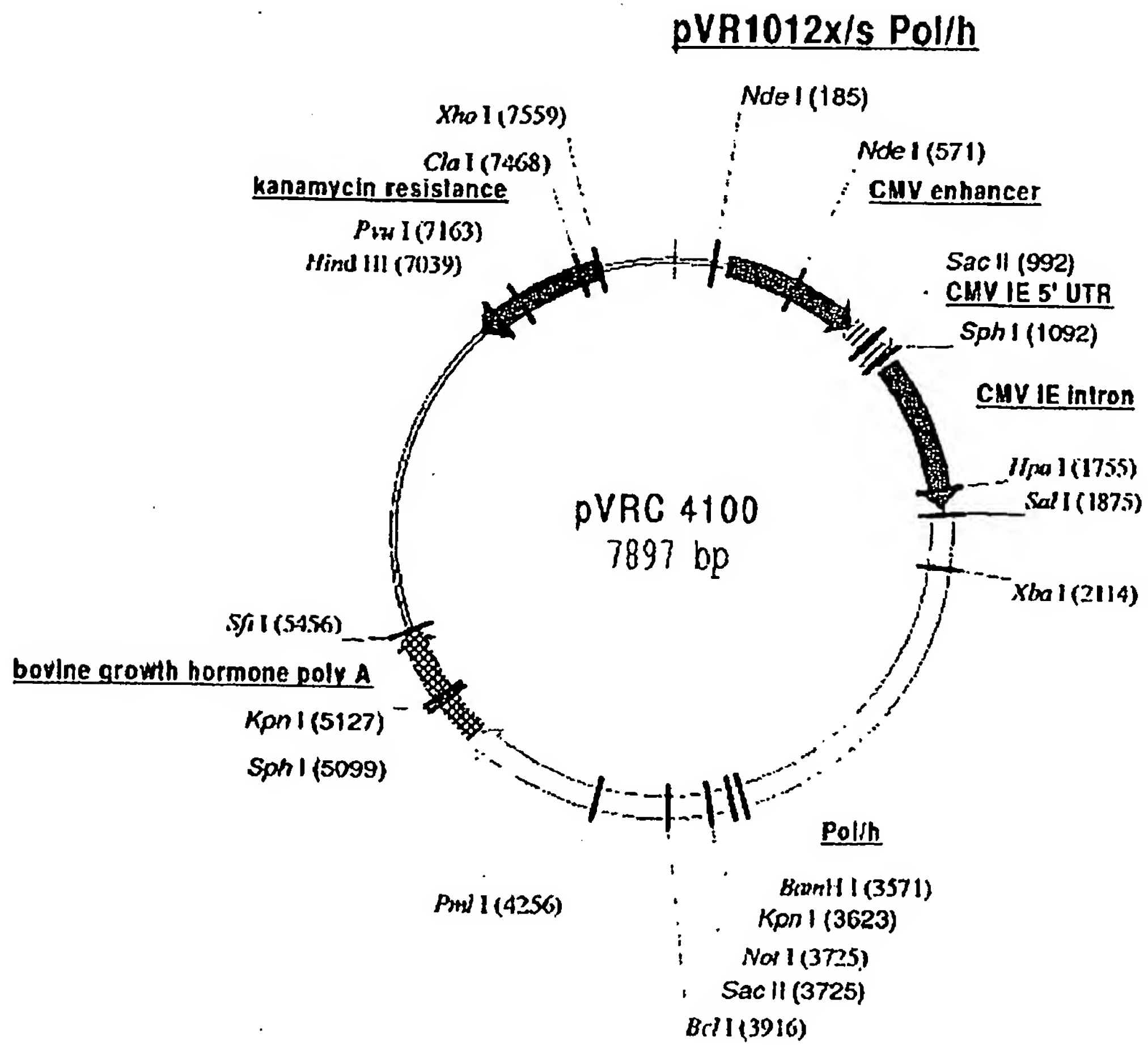


Figure 161

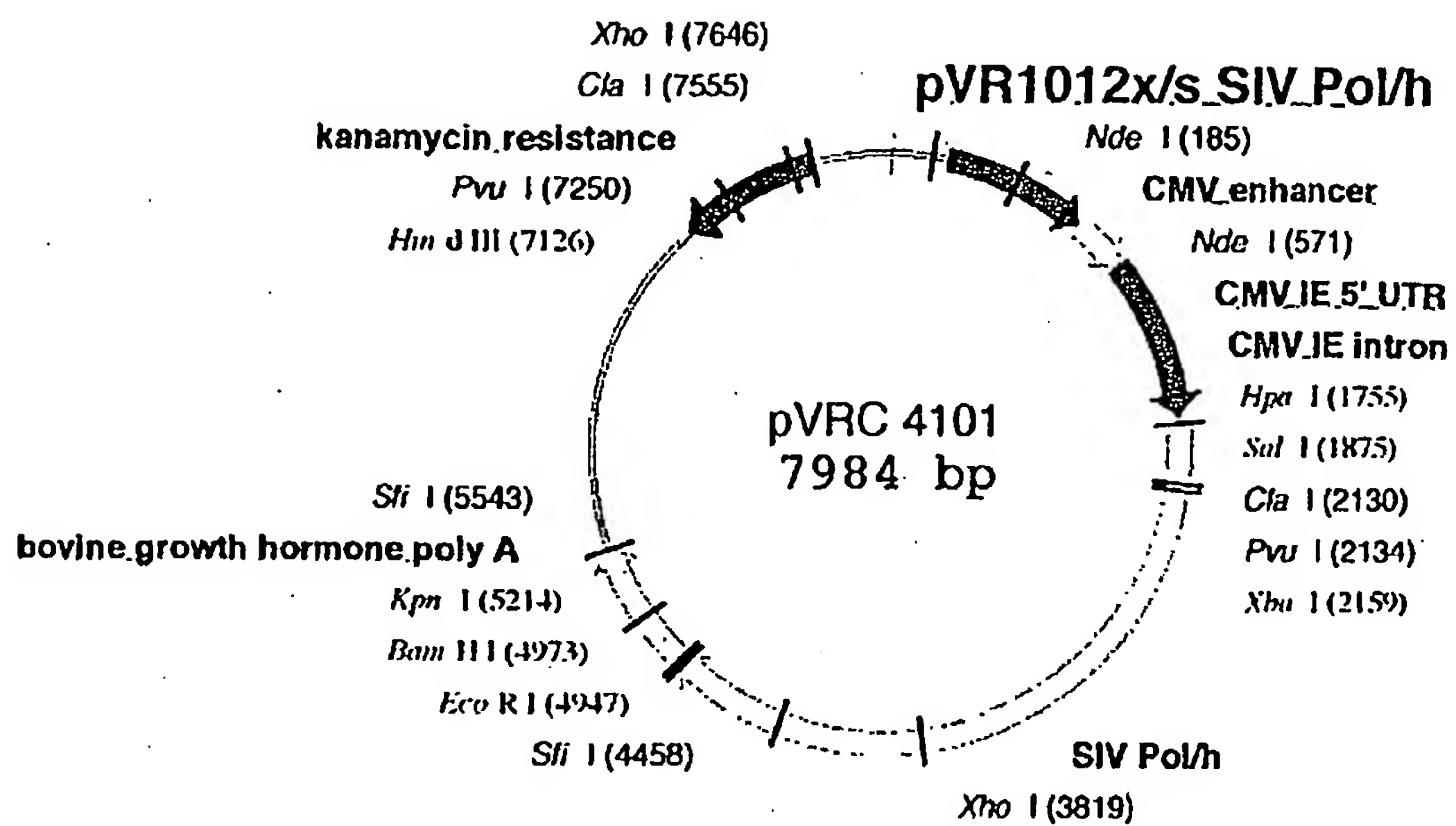


Figure 162

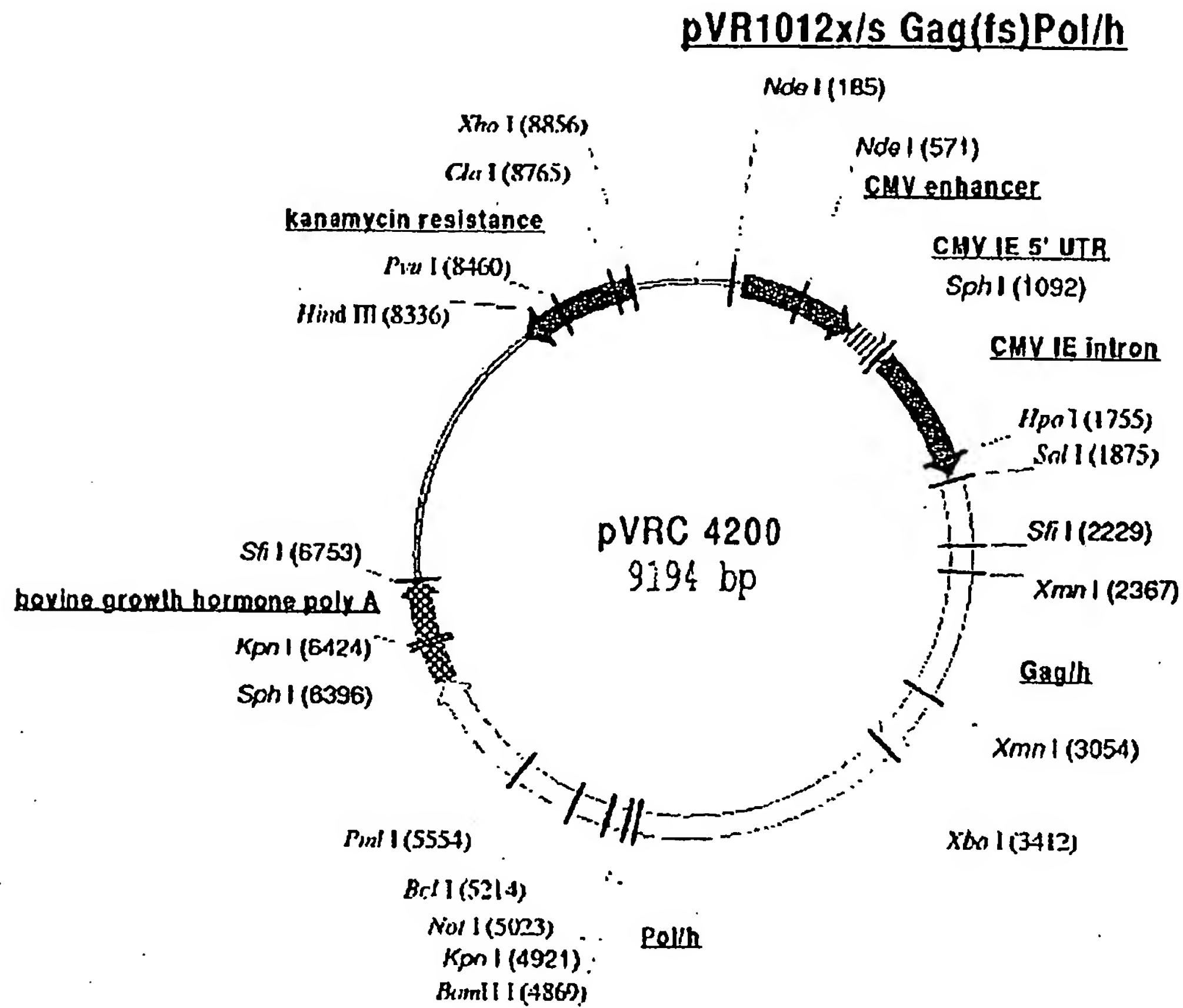


Figure 163

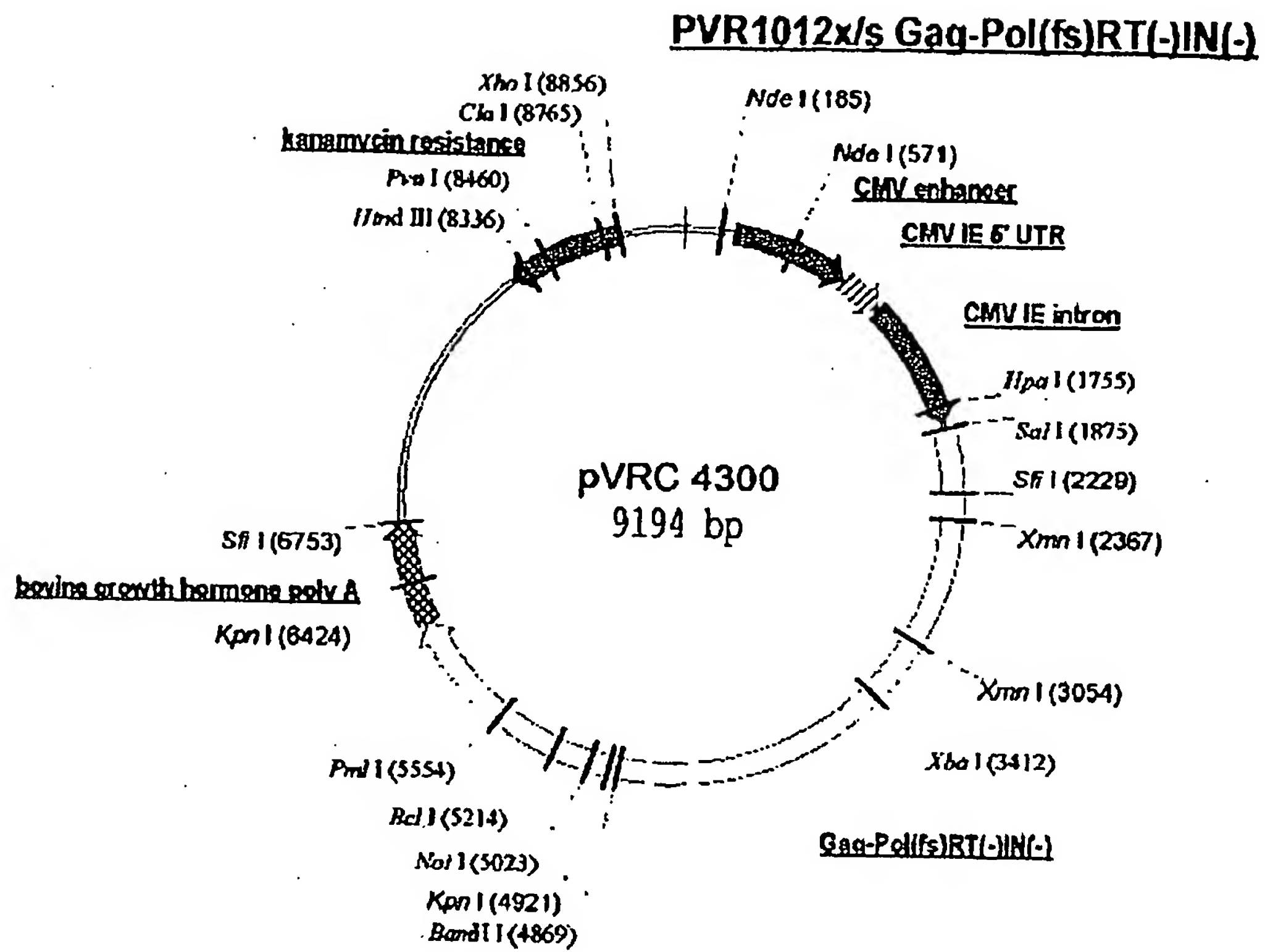


Figure 164

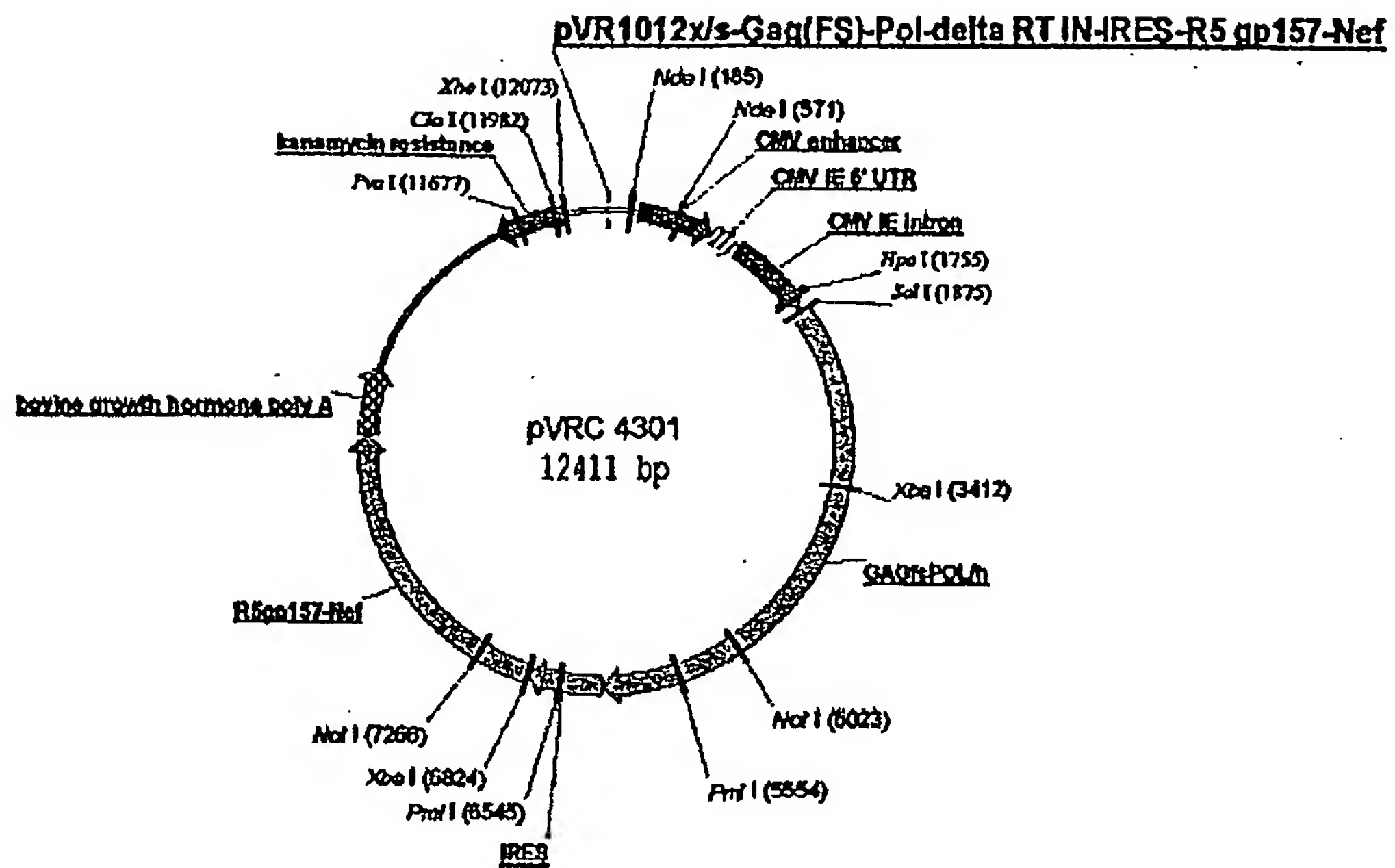


Figure 165

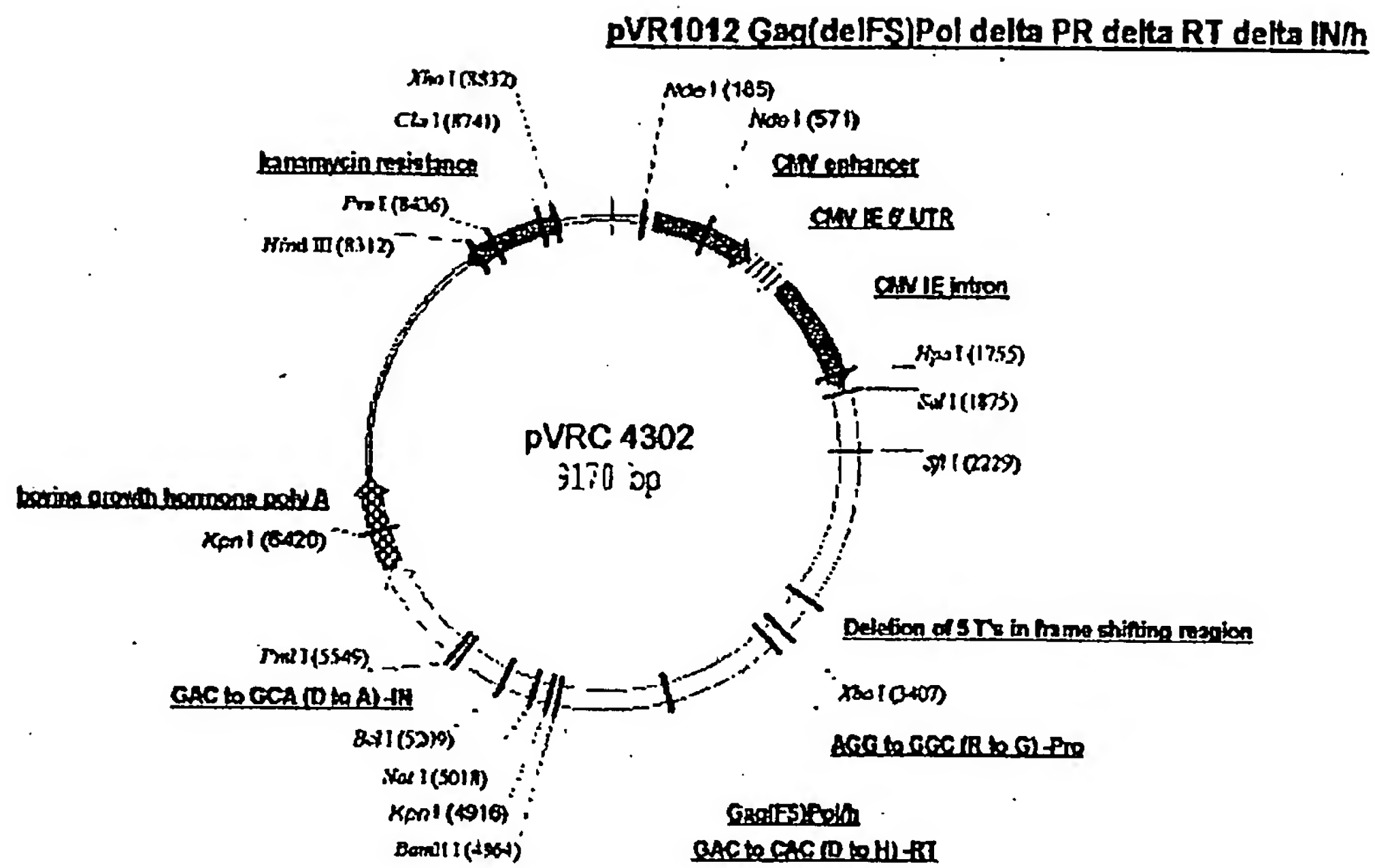


Figure 166

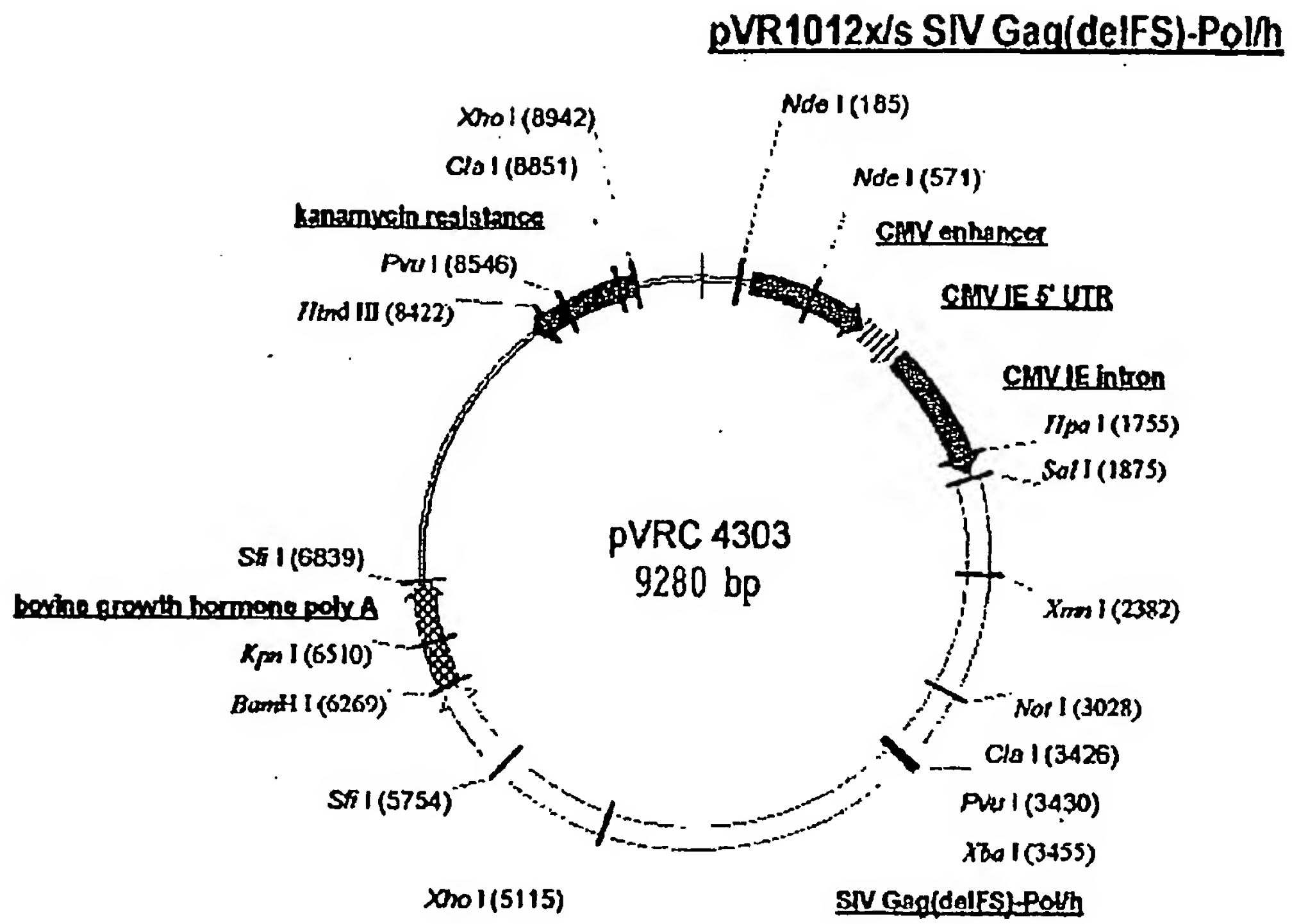


Figure 167



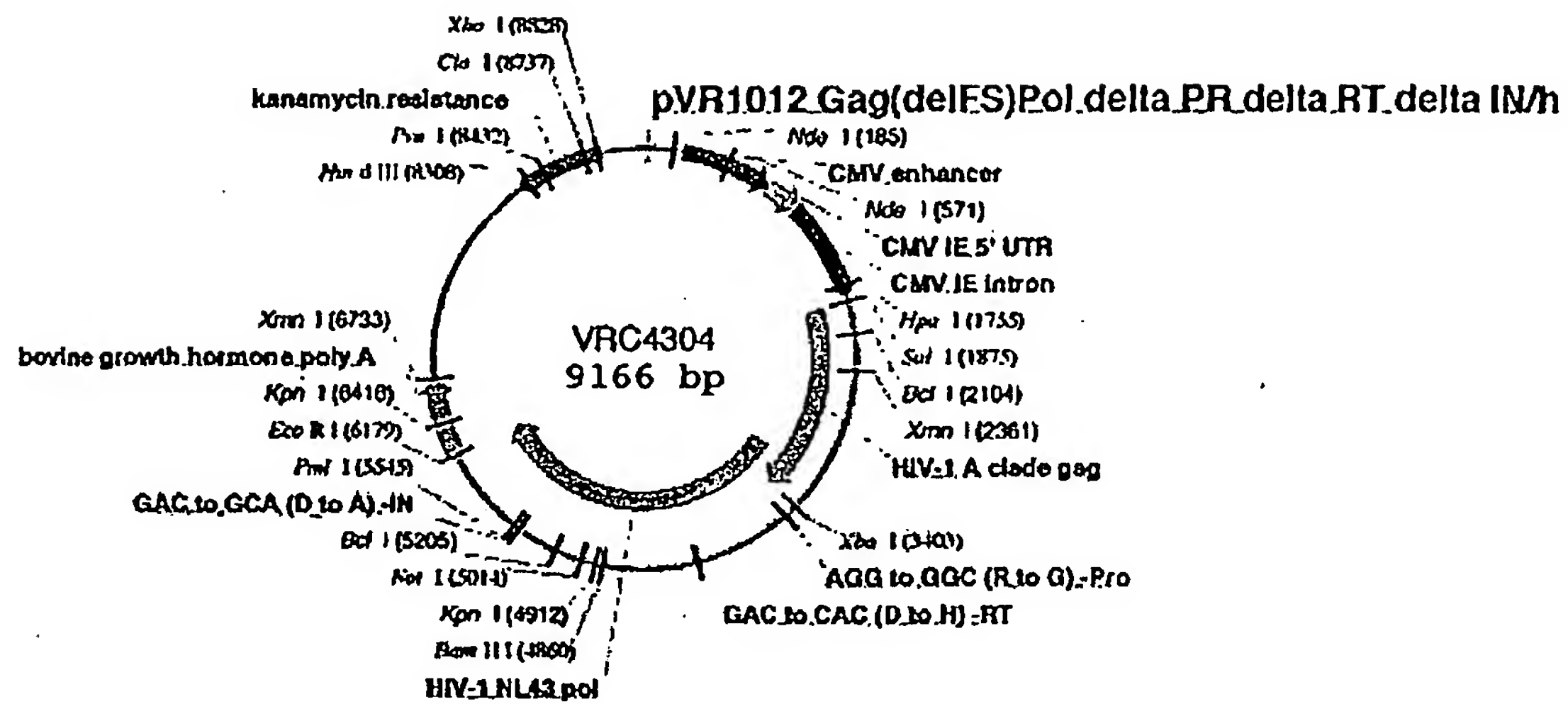


Figure 168

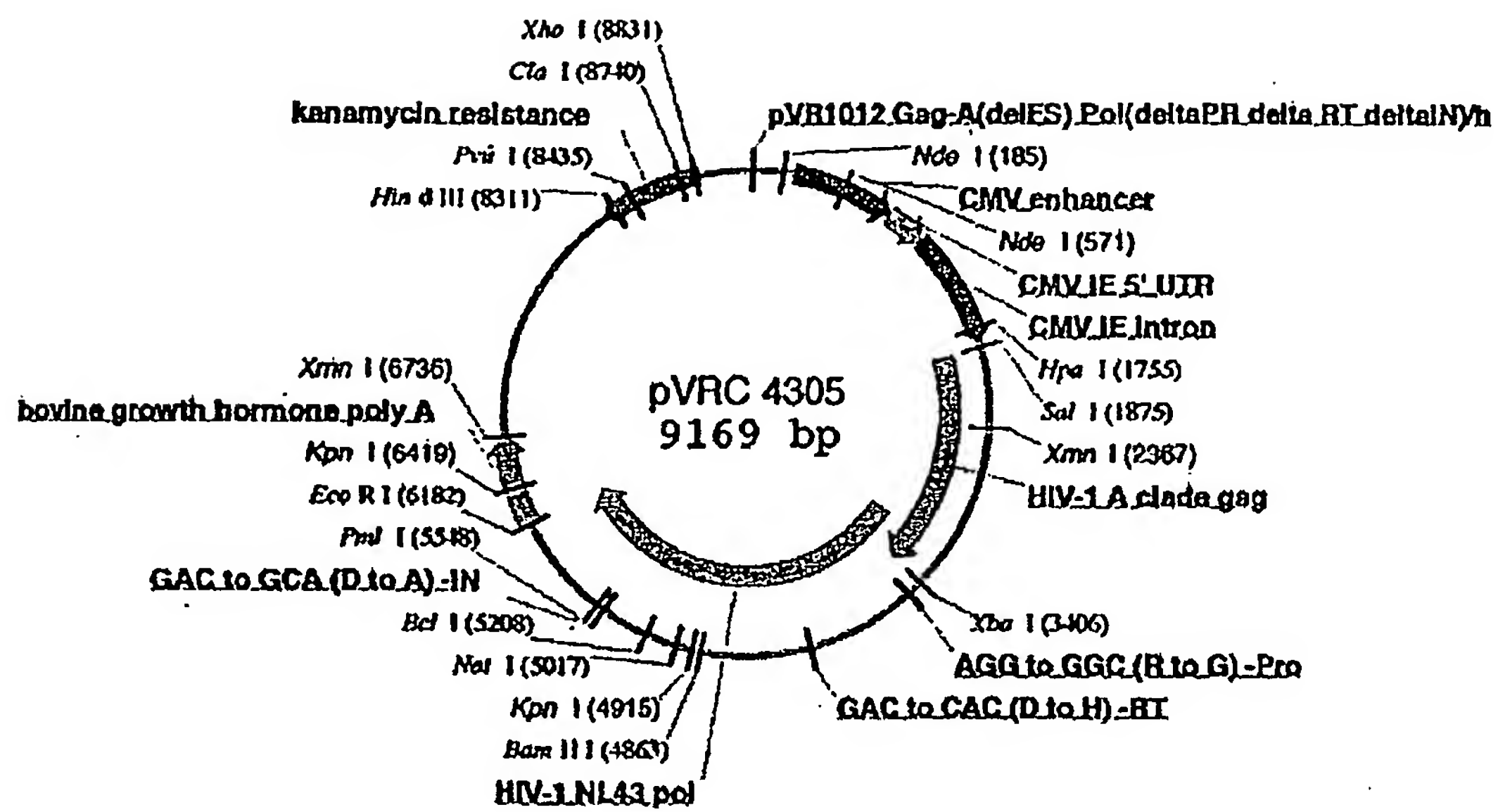


Figure 169

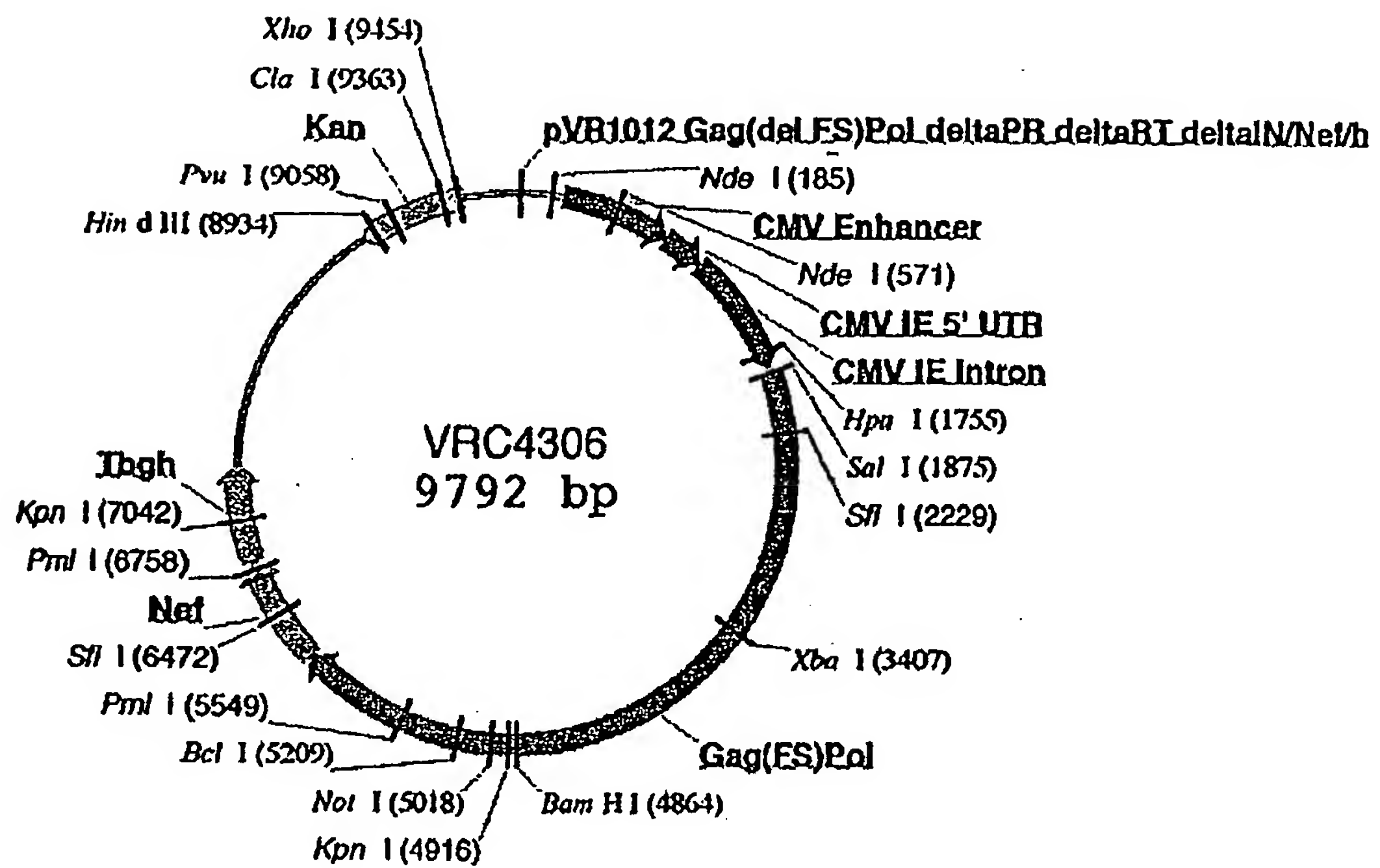


Figure 170

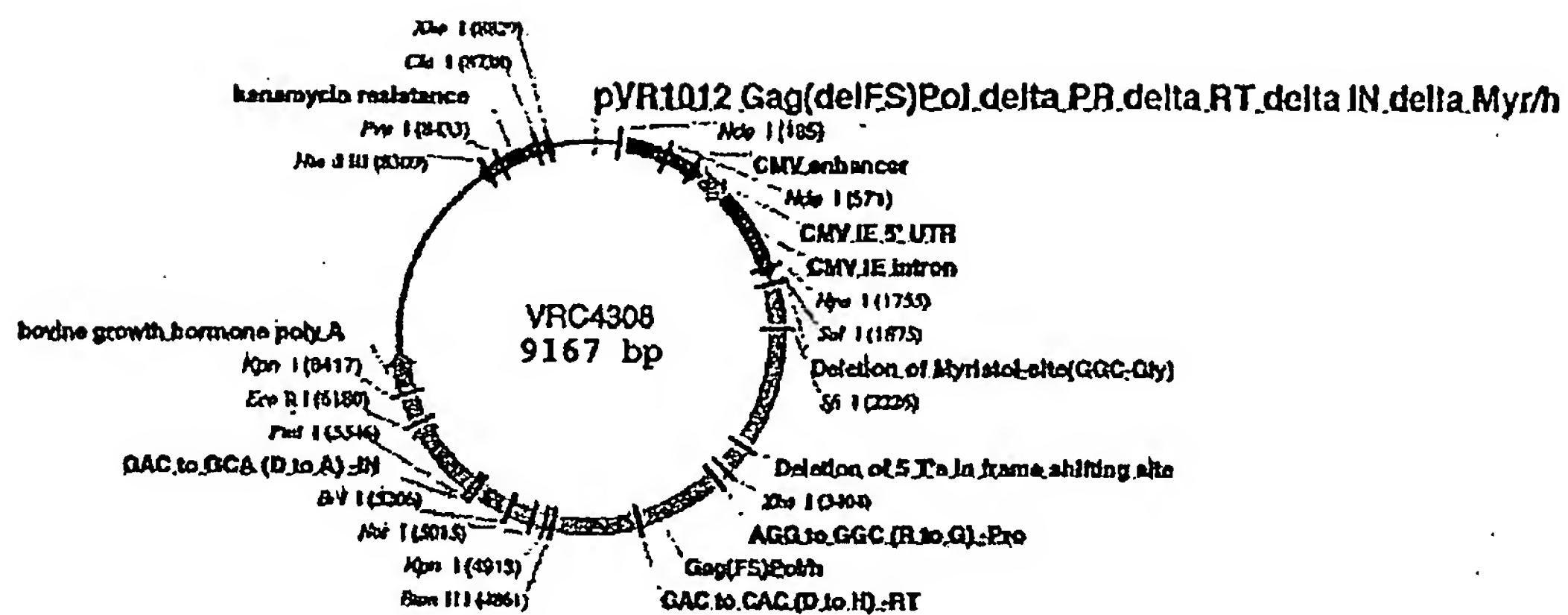


Figure 171

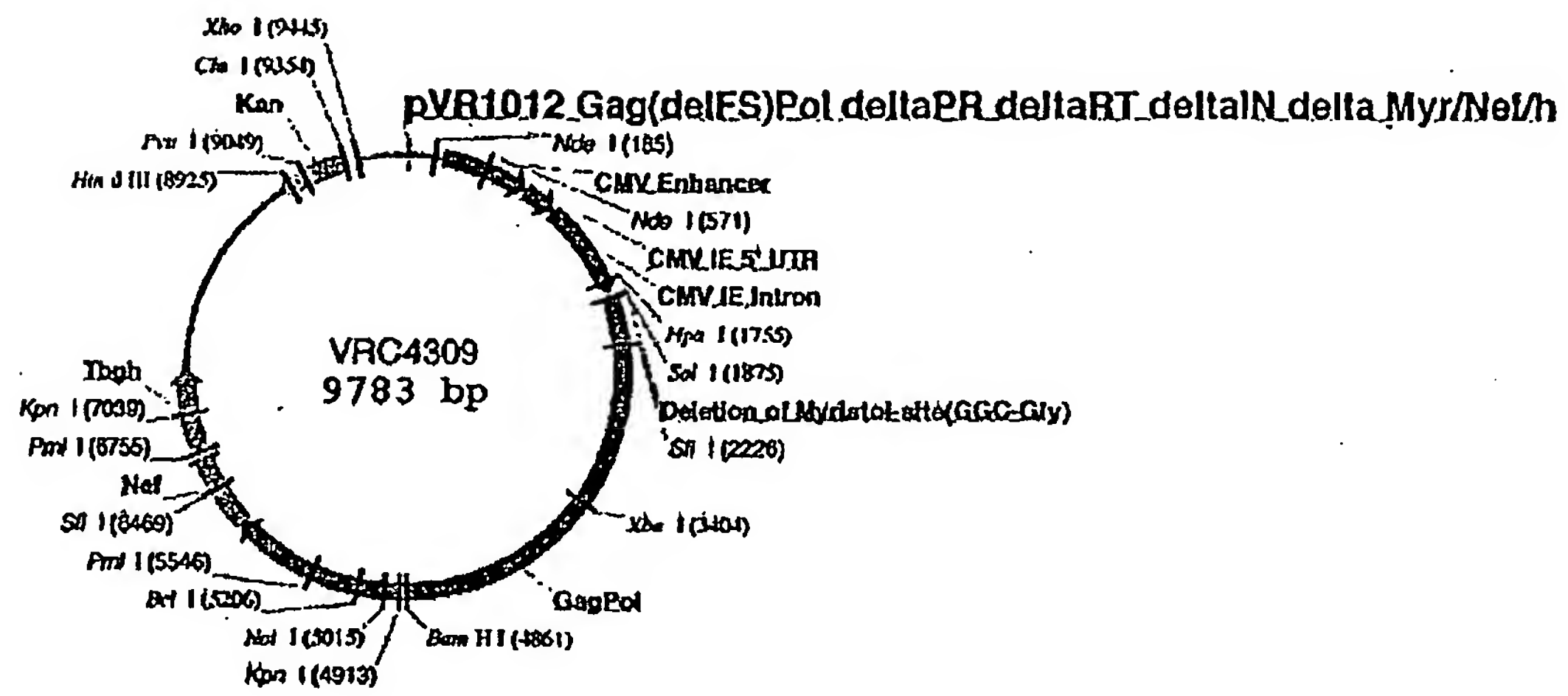


Figure 172

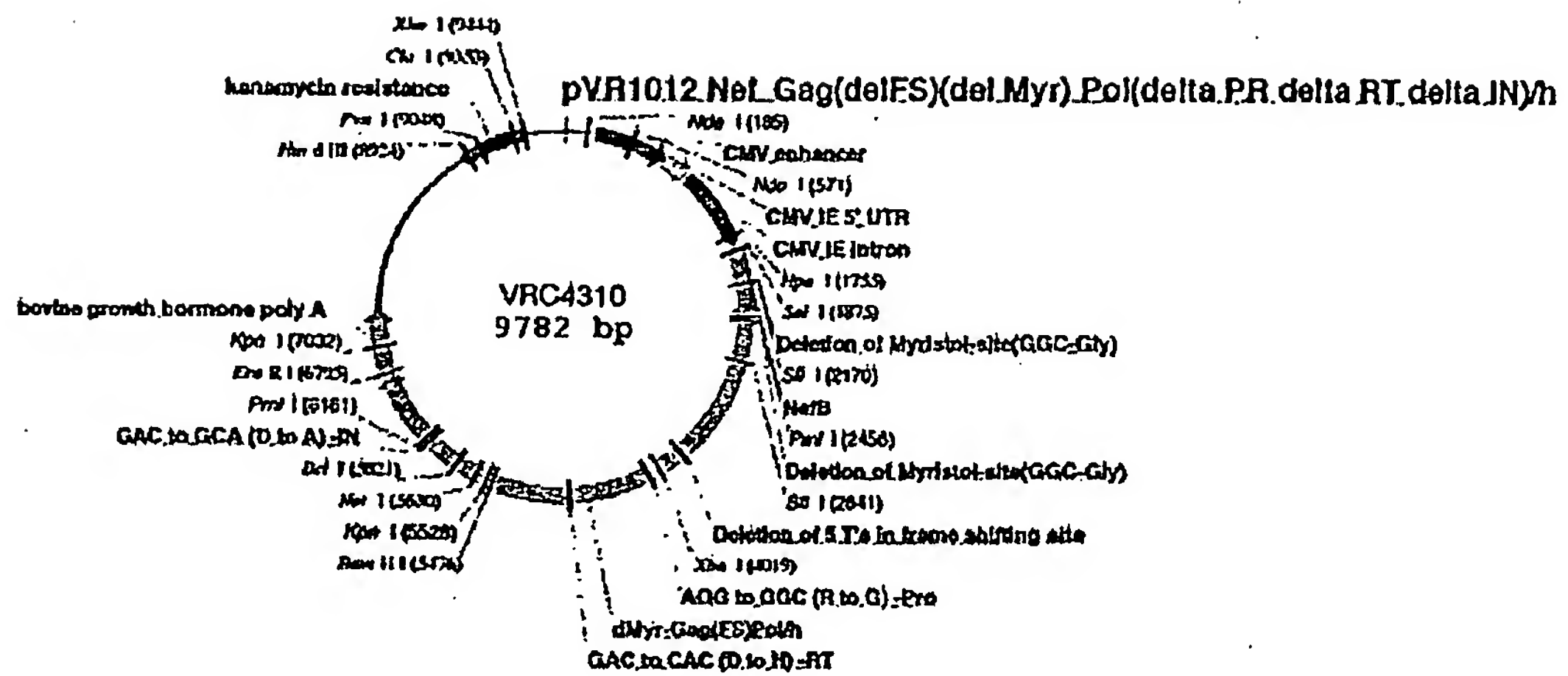


Figure 173

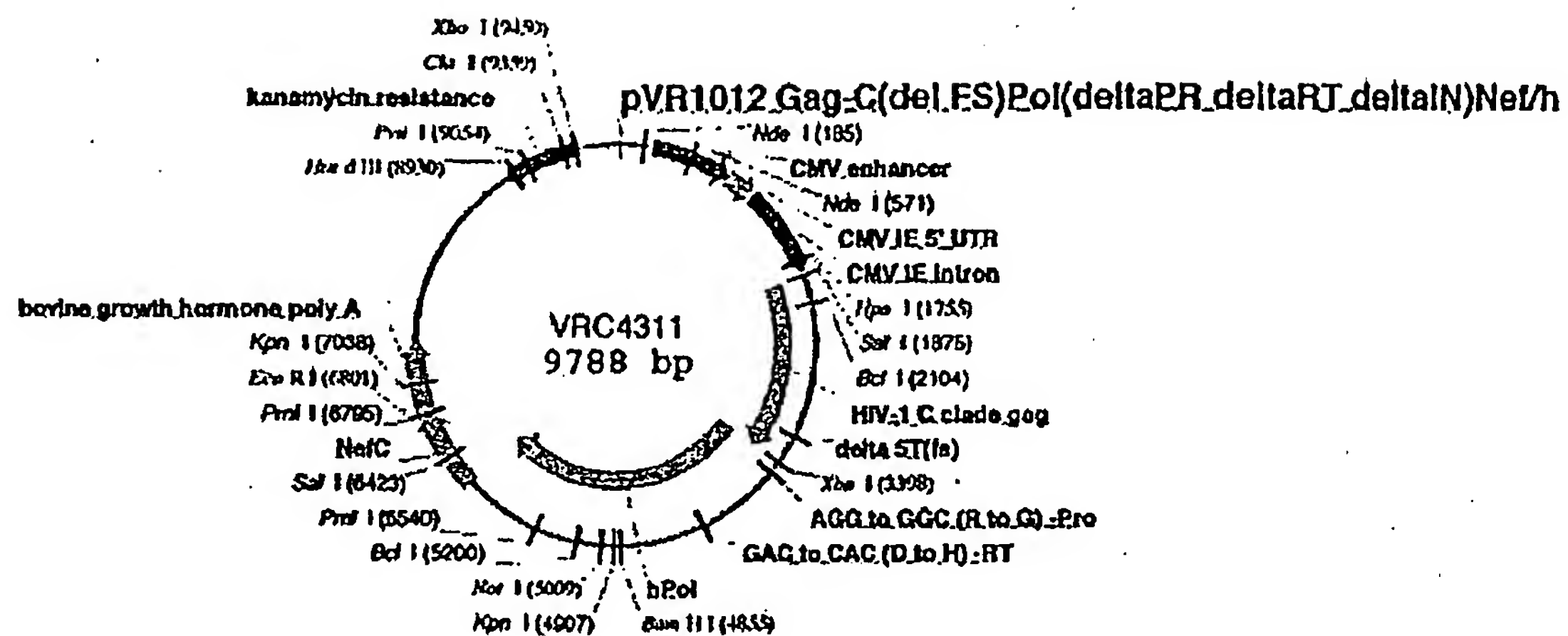


Figure 174

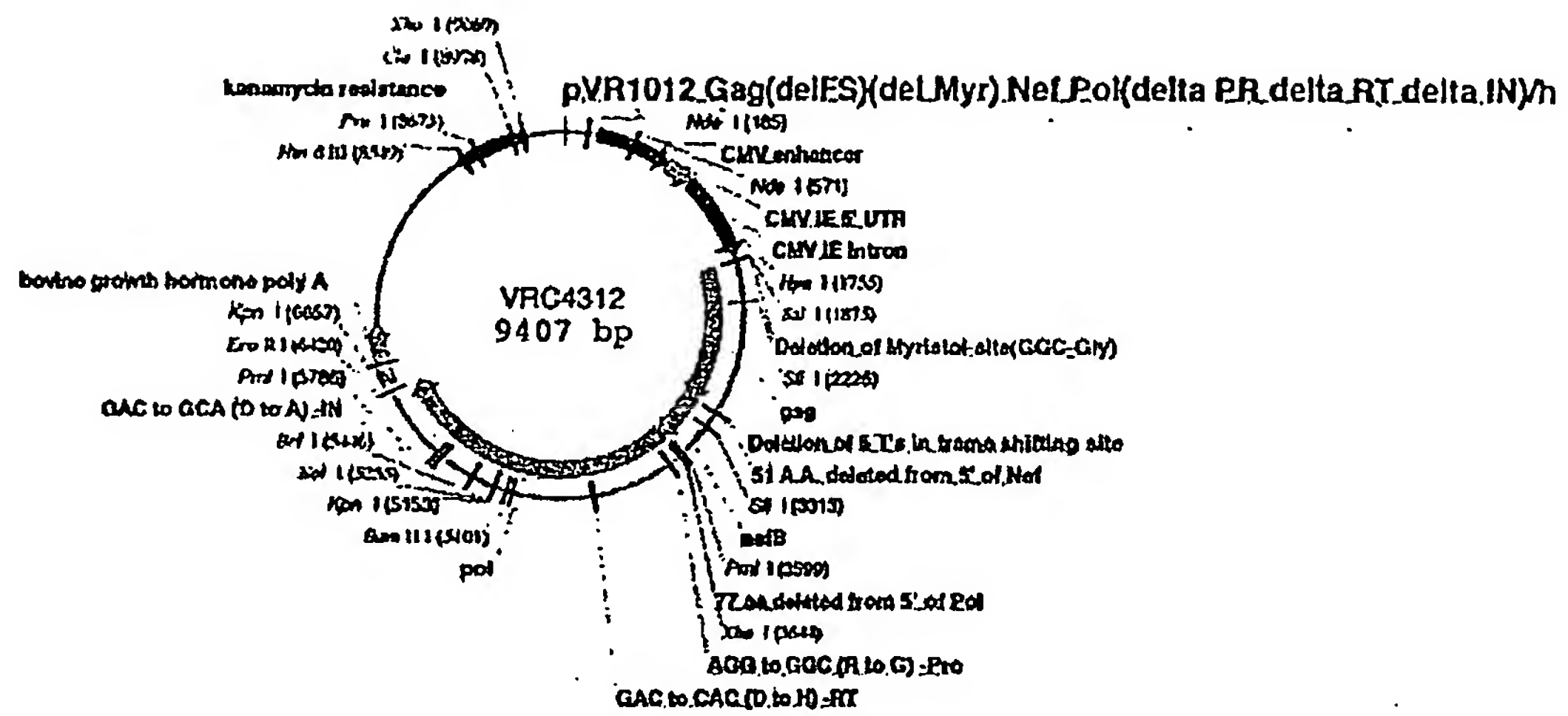


Figure 175



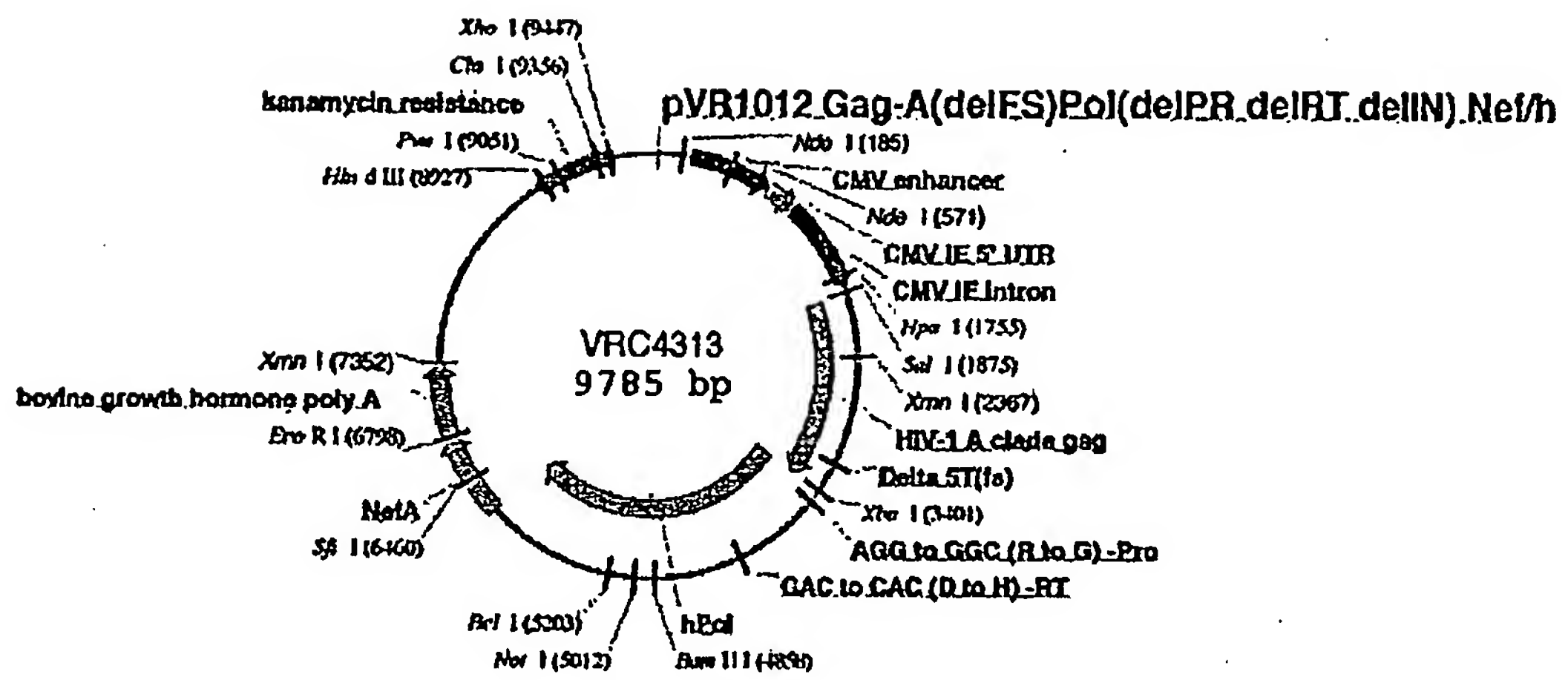


Figure 176

## SEQUENCE LISTING

<110> THE GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED  
BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES  
NABLE, Gary J.  
CHAKRABARTI, Bimal K.  
HUANG, Yue

<120> MODIFICATIONS OF HIV Env, Gag, AND Pol  
ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION

<130> NIH206.001VPC

<140> UNKNOWN

<141>

<150> US 60/225,097

<151> 2000-08-14

<150> US 60/252,115

<151> 2000-11-14

<150> US 60/279,257

<151> 2001-03-28

<160> 176

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 7547

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 1

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg accttcctac 660
ttggcagtac atctacgtat tagtcacgc taattaccatg gtgatgcggg ttgggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatgt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960

```

tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380  
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
agcggctcat ggctcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga acaccatgcg 1920  
cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980  
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040  
gcccgtgtgg aaggaggcca ccaccacct gctctgcgcc agcgacgcca aggcctacga 2100  
caccgaggtg cacaacgtgt gggccacca cgcctgcgtg cccaccgacc ccaaccccca 2160  
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220  
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280  
gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340  
cagcagcggc cgcatgatca tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400  
caccagcatc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460  
ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520  
ccaggcctgc cccaaggtga gcttcgagcc catccccaac cactactgcg ccccgccgg 2580  
cttcaccatc ctgaagtgca aggacaagaa gttcaacggc aaggggccct gcaccaacgt 2640  
gagcaccgtg cagtgcacc acggcatccg ccccggtgtg agcaccagc tgctgctgaa 2700  
cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc aacttcgccc acaacgcca 2760  
ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcacccgcc ccaacaaca 2820  
caccgcaag agcatccaca tcggccccgg ccgcgccttc tacaccaccg gcgagatcat 2880  
cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaaac acaccctgaa 2940  
caagatcgtg atcaagctgc gcgagcagtt cggcaacaag accatcgtgt tcaagcacag 3000  
cagcggcggc gaccccgaga tcgtgaccca cagcttcaac tgccggcggc agttcttcta 3060  
ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgagg 3120  
cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180  
caacatgtgg cagaaggtgg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240  
ctgcagcagc aacatcaccg gctgctgct gaccgcgac ggccgcaaca gcaacaacga 3300  
gagcgagatc ttccgcctgg gcggcgcgca catgcgcgac aactggcgca gcgagctgta 3360  
caagtacaag gtggtgaaga tcgagccct gggcgtggcc cccaccaagg ccaagcgccg 3420  
cgtggtgcag cgcgagaagc gcgccgtggg catcgccgcc ctgttccctgg gcttccctggg 3480  
cgccgcggc agcaccatgg gcgccgccag catgaccctg accgtgcagg cccgccagct 3540  
gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600  
cctgctgcag ctgaccgtgt ggggcatcaa gcagctgcag gcccgcatcc tggccgtgga 3660  
gcgctacctg aaggaccagc agctgctggg catctggggc tgcagcggca agcttatctg 3720  
caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780  
ccacaccacc tggatggagt gggaccgca gatcaacaac tacaccagcc tgatccacag 3840  
cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900  
caagtgggcc agcatgggcg gcaagtggag caagagcagc gtgatcggct ggcccgccgt 3960  
gcgcgagcgc atgcgccgcg ccgagcccg gcgcgacggc gtgggcgcgc tgagccgcga 4020  
cctggagaag cacggcgcca tcaccagcag caacaccgcc gccacaacg ccgcctgcgc 4080  
ctggctggag gcccaggagg aggaggaggt gggcttcgca gtgaccgcac aggtggcact 4140  
gcgcgcaatg acctacaagg ccgccgtgga cctgagccac ttcctgaagg agaaggcg 4200  
cctggagggc ctgatccaca gccagcggc ccaggacatc ctggacctgt ggatctacca 4260

cacccagggc	tacttccccg	actggcagaa	ctacaccccc	ggccccggcg	tgcgctaccc	4320
cctgaccttc	ggctggtgct	acaagctggt	gcccgtggag	cccgacaagg	tggaggaggc	4380
caacaagggc	gagaacacca	gcctgctgca	ccccgtgagc	ctgcacggca	tggcagcacc	4440
cgagcgcgag	gtgctggagt	ggcgcttcga	cagccgcctg	gccttccacc	acgtggcccc	4500
cgagctgcac	cccagtgact	tcaagaactg	ctgaggatcc	agatctgctg	tgccttctag	4560
ttgccagcca	tctgttgttt	gccccctccc	cgtgccttcc	btgaccctgg	aagggtgccac	4620
tcccactgtc	ctttcctaata	aaaatgagga	aattgcatcg	cattgtctga	gtaggtgtca	4680
ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	gaggattggg	aagacaatag	4740
caggcatgct	ggggatgcgg	tgggctctat	gggtacccag	gtgctgaaga	attgaccggg	4800
ttcctcctgg	gccagaaaga	agcaggcaca	tcccccttct	tgtgacacac	cctgtccacg	4860
cccctggttc	ttagttccag	ccccactcat	aggacactca	tagctcagga	gggctccgcc	4920
ttcaatccca	cccgtctaaag	tacttggagc	ggtctctccc	tccctcatca	gcccacccaaa	4980
ccaaacctag	cctccaagag	tgggaagaaa	ttaaagcaag	ataggctatt	aagtgcagag	5040
ggagagaaaa	tgcctccaac	atgtgaggaa	gtaatgagag	aatcataga	attttaaggc	5100
catcatggcc	ttaatcttcc	gcttcctcgc	tactgactc	gctgcgctcg	gtcgttcggc	5160
tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	5220
ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	5280
ccgcgttgct	ggcggtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	5340
gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	5400
gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	5460
ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	ctgtagggtat	ctcagttcgg	5520
tgtaggctgt	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	5580
gcgccttctc	cggtaactat	cgtcttgagt	ccaacccggt	aagacacgac	ttatcgccac	5640
tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	5700
tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	agtatttggt	atctgcgctc	5760
tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	5820
ccgctggtag	cgggtgggttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	5880
ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaaac	gaaaactcac	5940
gttaagggat	tttggtcattg	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	6000
aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttgggtct	gacagttacc	6060
aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttctgttc	tccatagttg	6120
cctgactcgg	gggggggggg	cgctgaggtc	tgcctcgtga	agaagggtgt	gctgactcat	6180
accaggcctg	aatcgcccca	tcattccagcc	agaaagtgag	ggagccacgg	ttgatgagag	6240
ctttgttgta	ggtggaccag	ttggtgattt	tgaacttttg	ctttgccacg	gaacggctctg	6300
cgttgctcgg	aagatgcgtg	atctgatect	tcaactcagc	aaaagtctga	tttattcaac	6360
aaagccgccg	tcccgtcaag	tcagcgtaat	gctctgccag	tgttacaacc	aattaaccaa	6420
ttctgattag	aaaaactcat	cgagcatcaa	atgaaactgc	aatttattca	tatcaggatt	6480
atcaatacca	tatttttgaa	aaagccgttt	ctgtaatgaa	ggagaaaact	caccgaggca	6540
gttccatagg	atggcaagat	cctggtatcg	gtctgcgatt	ccgactcgtc	caacatcaat	6600
acaacctatt	aatttccctt	cgtcaaaaat	aagggttatca	agtgagaaat	caccatgagt	6660
gacgactgaa	tccggtgaga	atggcaaaaag	cttatgcatt	tctttccaga	cttgttcaac	6720
aggccagcca	ttacgctcgt	catcaaaatc	actcgcatac	accaaaccgt	tattcattcg	6780
tgattgcgcc	tgagcgagac	gaaatacgcg	atcgctgtta	aaaggacaat	tacaaacagg	6840
aatcgaatgc	aaccggcgca	ggaacactgc	cagcgcatca	acaatatattt	cacctgaatc	6900
aggatattct	tctaatacct	ggaatgctgt	tttcccgggg	atcgcagtgg	tgagtaacca	6960
tgcattcatc	ggagtacgga	taaaatgctt	gatggtcgga	agaggcataa	attccgtcag	7020
ccagtttagt	ctgaccatct	catctgtaac	atcattggca	acgctacctt	tgccatgttt	7080
cagaaacaac	tctggcgcat	cgggcttccc	atacaatcga	tagattgtcg	cacctgattg	7140
cccgacatta	tgcgagagcc	atttataccc	atataaatca	gcattccatgt	tgggaatttaa	7200
tcgcggcctc	gagcaagacg	tttcccgttg	aatatggctc	ataacacccc	ttgtattact	7260
gtttatgtaa	gcagacagtt	ttattgttca	tgatgatata	tttttatctt	gtgcaatgta	7320
acatcagaga	ttttgagaca	caacgtggct	ttcccccccc	ccccattatt	gaagcattta	7380
tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	atttagaaaa	ataaacaat	7440
aggggttccg	cgcacatttc	cccgaaaagt	gccacctgac	gtctaagaaa	ccattattat	7500
catgacatta	acctataaaa	ataggcgat	cacgaggccc	tttcgtc		7547

&lt;210&gt; 2

&lt;211&gt; 8009

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 2

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgcgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agcctgggt ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcaag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgcgtccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga acaccatgcg 1920
cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgtggggca ccatgctgct 1980
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040
gcccggtgtg aaggaggcca ccaccacct gctctgcgcc agcgacgcca aggcctacga 2100
caccgaggtg cacaacgtgt gggccacca cgcctgcgtg ccaccgacc ccaaccccc 2160
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280
gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340
cagcagcggc cgcattgatc tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400
caccagcatc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460
ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520
ccaggcctgc cccaagggtg gcttcgagcc catcccaac cactactgcg ccccgccg 2580

```



cttcaccatc	ctgaagtgca	aggacaagaa	gttcaacggc	aagggcccct	gcaccaacgt	2640
gagcaccgtg	cagtgcaccc	acggcatccg	ccccgtggtg	agcaccacagc	tgctgctgaa	2700
cgaccgcctg	gccgaggagg	aggtggtgat	ccgcagcgcc	aacttcgccg	acaacgcca	2760
ggtgatcatc	gtgcagctga	acgagagcgt	ggagatcaac	tgaccccgcc	ccaacaacaa	2820
cacccgcaag	agcatccaca	tccgcccccg	ccgcgccttc	tacaccaccg	gcgagatcat	2880
cggcgacatc	cgccaggccc	actgcaacct	gagccgcgcc	aagtggaaacg	acaccctgaa	2940
caagatcgtg	atcaagctgc	gcgagcagtt	cggcaacaag	accatcgtgt	tcaagcacag	3000
cagcggcgcc	gaccccgaga	tcgtgaccca	cagcttcaac	tgccggcgccg	agttcttcta	3060
ctgcaacagc	acccagctgt	tcaacagcac	ctggttcaac	agcacctgga	gcaccgaggg	3120
cagcaacaac	accgagggca	gcgacaccat	caccctgccc	tgccgcataca	agcagatcat	3180
caacatgttg	cagaagggtg	gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	3240
ctgcagcagc	aacatcaccg	gcctgctgct	gacccgcgac	ggcggcaaca	gcaacaacga	3300
gagcgagatc	ttccgcctgg	gcggcgggcg	catgcgcgac	aactggcgca	gcgagctgta	3360
caagtacaag	gtggtgaaga	tcgagccctt	gggcgtggcc	cccaccaagg	ccaagcgccg	3420
cgtggtgcag	cgcgagaagc	gcgccgtggg	catcggcgcc	ctgttccctg	gcttccctgg	3480
cgcgcgcggc	agcaccatgg	gcgccgccag	catgacctg	accgtgcagg	cccgccagct	3540
gctgagcggc	atcgtgcagc	agcagaacaa	cctgctgcgc	gccatcgagg	cccagcagca	3600
cctgctgcag	ctgacctgtg	ggggcatcaa	gcagctgcag	gcccgcatacc	tgcccggtga	3660
gcgctacctg	aaggaccagc	agctgctggg	catctggggc	tgacgcggca	agcttatctg	3720
caccaccgcc	gtgccctgga	acgccagctg	gagcaacaag	agcctggagc	agatctggaa	3780
ccacaccacc	tggatggagt	gggaccgcga	gatcaacaac	tacaccagcc	tgatccacag	3840
cctgatcgag	gagagccaga	accagcagga	gaagaacgag	caggagctgc	tggagctgga	3900
caagtgggcc	agcctgtgga	actggttcaa	catcaccaac	tggctgtggt	acatcaagct	3960
gttcatcatg	atcgtgggcg	gcctggtggg	cctgcgcata	gtgttcgccg	tgctgagcat	4020
cgtgaaccgc	gtgcgccagg	gctacagccc	cctgagcttc	cagacccacc	tgcccatccc	4080
ccgcggcccc	gaccgccccg	agggcatcga	ggaggagggg	ggcgagcgcg	accgcgaccg	4140
cacgatccgc	ctggtgaacg	gcagcctggc	cctgatctgg	gacgacctgc	gcagcctgtg	4200
cctgttcagc	taccaccgcc	tgccgcacct	gctgctgate	gtgaccgcga	tcgtggagct	4260
gctgggcccgc	cgcggctggg	aggccctgaa	gtactggtgg	aacctgctgc	agtactggag	4320
ccaggagctg	aagaacagcg	ccgtgagcct	gctgaacgcc	accgccatca	tgggcggcaa	4380
gtggagcaag	agcagcgtga	tcggctggcc	cgccgtgcgc	gagcgcatgc	gccgcgccga	4440
gcccgcgcgc	gacggcgtgg	gcgccgtgag	ccgcgacctg	gagaagcacg	gcgccatcac	4500
cagcagcaac	accgccgcca	acaacgccgc	ctgcgcctgg	ctggaggccc	aggaggagga	4560
ggagggtggc	ttcgcagtg	ccgcacaggt	ggcactgcgc	gcaatgacct	acaaggccgc	4620
cgtggacctg	agccacttcc	tgaaggagaa	gggcggcctg	gagggcctga	tccacagcca	4680
gcgccgccag	gacatcctgg	acctgtggat	ctaccacacc	cagggctact	tccccgactg	4740
gcagaactac	acccccggcc	ccggcgtgcg	ctacccccctg	accttcggct	ggtgctacaa	4800
gctggtgccc	gtggagcccc	acaagggtga	ggaggccaac	aagggcgaga	acaccagcct	4860
gctgcacccc	gtgagcctgc	acggcatggc	agcacccgag	cgcgaggtgc	tggagtggcg	4920
cttcgacagc	cgcctggcct	tccaccacgt	ggcccgcgag	ctgcaccccg	agtacttcaa	4980
gaactgctga	ggatccggat	ccagatctgc	tgtgccttct	agttgccagc	catctgttgt	5040
ttgccccctc	cccgtgcctt	ccttgaccct	ggaagggtgc	actcccactg	tcctttccta	5100
ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	tgggggggtg	5160
ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	5220
ggtgggctct	atgggtaccc	aggtgctgaa	gaattgaccc	ggttctctct	gggccagaaa	5280
gaagcaggca	catcccttcc	tctgtgacac	acctgttcca	cgcctctggt	tcttagttcc	5340
agccccactc	ataggacact	catagctcag	gagggctccg	ccttcaatcc	caccgcctaa	5400
agtacttgga	gcggtctctc	cctccctcat	cagccccacca	aaccaaacct	agcctccaag	5460
agtgggaaga	aattaaagca	agataggcta	ttaagtgcag	aggagagaaa	aatgcctcca	5520
acatgtgagg	aagtaatgag	agaaatcata	gaattttaag	gcatcatggg	ccttaatctt	5580
ccgcttcttc	gctcactgac	tcgctgcgct	cggtcgttct	gctgcggcga	gcggtatcag	5640
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	5700
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgcttt	5760
tccataggct	ccgccccctt	gacgagcatc	acaaaaatcg	acgctcaagt	cagagggtggc	5820
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaaagctcc	ctcgtgcgct	5880

```

ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 5940
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 6000
agctgggctg tgtgcacgaa cccccggtc agcccgaccg ctgcgcctta tccggtaact 6060
atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca gccactggta 6120
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta 6180
actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 6240
tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggg agcgggtgggt 6300
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 6360
tcttttctac ggggtctgac gctcagtggg acgaaaactc acgttaaggg attttggtca 6420
tgagattatc aaaaaggatc ttcacctaga tctttttaa ttaaaaatga agttttaaat 6480
caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg 6540
cacctatctc agcgatctgt ctatttcggt catccatagt tgcctgactc gggggggggg 6600
ggcgctgagg tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc 6660
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 6720
agttgggtgat tttgaacttt tgctttgccg cggaacgggt tgcgtgtcgc ggaagatgcg 6780
tgatctgata cttcaactca gcaaaagttc gatatttca acaaagccgc cgtcccgta 6840
agtcagcgta atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc 6900
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 6960
aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 7020
atcctgggat cgggtctgca ttccgactcg tccaacatca atacaacct ttaattttcc 7080
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccgggtg 7140
gaatggcaaa agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc 7200
gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag 7260
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg 7320
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 7380
ctggaatgct gttttcccg ggatcgcat ggtgagtaac catgcatcat caggagtacg 7440
gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgacct 7500
ctcatctgta acatcattgg caacgctacc tttgccaatg ttcagaaaca actctggcgc 7560
atcgggcttc ccatacaatc gatagattgt cgcacctgat tgcccagacat tatcgcgagc 7620
ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 7680
cgtttcccg tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag 7740
ttttattgtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 7800
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 7860
gagcggatac atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt 7920
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 7980
aaataggcgt atcacgaggc ctttctgctc

```

8009

&lt;210&gt; 3

&lt;211&gt; 7558

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 3

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420

```

```

ccccctggc tgaccgccc accgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcatatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtag atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcggtgacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcccg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg ggccctatac acccccgtct ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcggtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagc gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 2160
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaaggtgag ctctcgagccc atccccacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggccccctg caccaacgtg 2640
agcaccgtgc agtgacacca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcaccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac categgcaag 2880
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagacacca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcacc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cageggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctggggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctggggc tggcccccac caaggccaag 3420
cgccgcgtgg tgcagcgca gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480
ctgggcccgc ccggcagcac catgggcgcc gccagcatga cctgaccgt gcaggcccgc 3540
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720

```



atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780  
 tggaaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgac 3840  
 cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900  
 ctggacaagt gggccagcat gggcggcaag tggagcaaga gcagcgtgat cggctggccc 3960  
 gccgtgcgcg agcgcctgcg ccgcgccgag cccgccgccg acggcgtggg cggcgtgagc 4020  
 cgcgacctgg agaagcacgg cgccatcacc agcagcaaca ccgccgcca caacgccgcc 4080  
 tgcgcctggc tggaggccca ggaggaggag gaggtgggct tcgcagtac cgcacagggtg 4140  
 gcactgcgcg caatgacct caaggccgcc gtggacctga gccacttctt gaaggagaag 4200  
 ggccgcctgg agggcctgat ccacagccag cgccgccagg acatcctgga cctgtggatc 4260  
 taccacaccc agggctactt ccccgactgg cagaactaca ccccgggccc cggcgtgcgc 4320  
 taccctctga ccttcggctg gtgtacaag ctggtgcccg tggagcccga caagggtggag 4380  
 gaggcccaaca agggcgagaa caccagcctg ctgcaccccg tgagcctgca cggcatggca 4440  
 gcacccgagc gcgaggtgct ggagtggcgc ttcgacagcc gcctggcctt ccaccacgtg 4500  
 gcccgcgagc tgcaccccga gtacttcaag aactgctgag gatccggatc cagatctgct 4560  
 gtgccttcta gttgccagcc atctgttgtt tgcctctccc ccgtgccttc cttgacctg 4620  
 gaaggtgcca ctcccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg 4680  
 agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg 4740  
 gaagacaata gcaggcatgc tggggatgcg gtgggctcta tgggtacca ggtgctgaag 4800  
 aattgacccg gttcctcctg ggccagaaag aagcaggcac atccccttct ctgtgacaca 4860  
 cctgtccac gccctgggtt cttagtcca gccccactca taggacactc atagctcagg 4920  
 agggctccgc cttcaatccc acccgctaaa gtacttgagg cggctctctc ctcctcatc 4980  
 agccaccaa accaaaccta gcctccaaga gtgggaagaa attaaagcaa gataggctat 5040  
 taagtgcaga gggagagaaa atgcctccaa catgtgagga agtaatgaga gaaatcatag 5100  
 aattttaagg ccatcatggc cttaatcttc cgcttcctcg ctactgact cgtgcgctc 5160  
 ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtaatatc ggttatccac 5220  
 agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 5280  
 ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc cgccccctg acgagcatca 5340  
 caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 5400  
 gtttccccct ggaagctccc tcgtgcgctc tctgttccg accctgccgc ttaccggata 5460  
 cctgtccgcc tttctccctt cgggaagcgt ggcgtttct catagctcac gctgtaggta 5520  
 tctcagttcg gtgtaggctg ttcgtccaa gctgggctgt gtgcacgaac ccccggttca 5580  
 gcccgaccgc tgcgccttat ccggtacta tcgtcttgag tccaacccgg taagacacga 5640  
 cttatcgcca ctggcagcag ccaactgtaa caggattagc agagcgaggt atgtaggcgg 5700  
 tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 5760  
 tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 5820  
 caaacaacc accgctggta gcggtggttt ttttggttgc aagcagcaga ttacgcgcag 5880  
 aaaaaaagga tctcaagaag atcctttgat cttttctacg ggggtctgacg ctcagtggaa 5940  
 cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 6000  
 ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 6060  
 tgacagttac caatgcttaa tcagtaggc acctatctca gcgatctgtc tatttcgttc 6120  
 atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 6180  
 tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 6240  
 gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 6300  
 ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 6360  
 atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 6420  
 caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttattc 6480  
 atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 6540  
 tcaccgaggc agttccatag gatggcaaga tcctggatatc ggtctgcgat tccgactcgt 6600  
 ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 6660  
 tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 6720  
 acttggtcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaacg 6780  
 ttattcatte gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6840  
 ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6900  
 tcacctgaat caggatatte ttctaatacc tggaatgctg ttttccggg gatcgagtg 6960  
 gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggctcg aagaggcata 7020

```

aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 7080
ttgccatggt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 7140
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 7200
ttggaattta atcgcggcct cgagcaagac gtttcccggt gaatatggct cataacaccc 7260
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 7320
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc cccccattat 7380
tgaagcattt atcaggggta ttgtctcatg agcggatata tatttgaatg tatttagaaa 7440
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 7500
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 7558

```

&lt;210&gt; 4

&lt;211&gt; 7351

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 4

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggtctg gggcctatac acccccgtct ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcctg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcaatcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040

```

```

cccggtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgccaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcgggc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaaggtgag ctctgagccc atccccaacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggccccctg caccaacgtg 2640
agcacccgtg agtgcaccca cggcatccgc ccggtggtga gcaccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaa 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcaccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggccccac caaggccaag 3420
cgccgcgtgg tgcagcgcgga gaagcgcgcc gtgggcacat gcgccctgtt cctgggcttc 3480
ctgggcgccc ccggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccg 3540
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgaggcccc catcctggcc 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
ggtggcaagt ggtcaaaaag tagtgtgatt ggatggcctg ctgtaaggga aagaatgaga 3780
cgagctgagc cagcagcaga tgggggtggga gcagtatctc gagacctaga aaaacatgga 3840
gcaatcacia gtagcaatac agcagctaac aatgctgctt gtgcctggct agaagcacia 3900
gaggaggaag aggtgggttt tccagtcaca cctcaggtac ctttaagacc aatgacttac 3960
aaggcagctg tagatcttag ccacttttta aaagaaaagg ggggactgga agggctaatt 4020
cactcccaaa gaagacaaga tacccttgat ctgtggatct accacacaca aggctacttc 4080
cctgattggc agaactacac accaggggcca ggggtcagat atccactgac ctttggatgg 4140
tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccataa aggagagaac 4200
accagcttgt tacaccctgt gagcctgcat ggaatggatg accctgagag agaagtgtta 4260
gagtggaggt ttgacagccg cctagcatct catcacgtgg cccgagagct gcacccggag 4320
tacttcaaga actgctgagg atccagatct gctgtgcctt ctagtbtgca gccatctgtt 4380
gtttgcccc tccccgtgcc ttccttgacc ctggaagggt ccactccac tgtcctttcc 4440
taataaaatg aggaaattgc atcgcatgtt ctgagtaggt gtcattctat tctgggggg 4500
ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat 4560
gcggtgggct ctatgggtac ccaggtgctg aagaattgac ccggttcctc ctgggccaga 4620
aagaagcagg cacatccctt tctctgtgac acaccctgtc cagccccctg gttcttagtt 4680
ccagccccac tcataggaca ctcatagctc aggagggtc cgccctcaat cccaccgct 4740
aaagtacttg gagcggcttc tccctccctc atcagcccac caaaccaaac ctagcctcca 4800
agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc 4860
caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatcat ggccttaatc 4920
ttccgcttcc tcgctcactg actcgctgcg ctcggtcggt cggctgcggc gagcgggtatc 4980
agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa 5040
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgctg tgctggcgctt 5100
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 5160
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgctgcg 5220
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 5280
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 5340

```



```

caagctgggc tgtgtgcacg aaccccccg tccagcccgac cgctgcgcct tatccggtaa 5400
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 5460
taacaggatt agcagagcga ggtatgtagg cgggtgtaca gagttcttga agtgggtggc 5520
taactacggc tacactagaa gaacagtatt tggatatctgc gctctgctga agccagttac 5580
cttcggaaaa agagttggta gctcttgatc cggcaaaaca accaccgctg gtagcgggtg 5640
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5700
gatcttttct acgggggtctg acgctcagtg gaacgaaaac tcacgttaag ggatttttgg 5760
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttta 5820
atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 5880
ggcacctatc tcagcgatct gtctatttctg ttcattccata gttgcctgac tcgggggggg 5940
ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 6000
cccatcatcc agccagaaaag tgagggagcc acggttgatg agagctttgt tgtaggtgga 6060
ccagttgggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg 6120
cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccg 6180
caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac 6240
tcacgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatbtt 6300
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 6360
agatcctggg atcgggtctgc gattccgact cgtccaacat caatacaacc tattaatttc 6420
ccctcgtcaa aaataagggt atcaagttag aaatcaccat gagtgacgac tgaatccgg 6480
gagaatggca aaagcttatg catttctttc cagacttggt caacaggcca gccattacgc 6540
tcgtcatcaa aatcactcgc atcaaccaa cgttattca ttcgtgattg cgcctgagcg 6600
agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg 6660
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttcta 6720
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcac atcaggagta 6780
cggataaaat gcttgatggg cgggaagggc ataaattccg tcagccagtt tagtctgacc 6840
atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc 6900
gcacggggct tcccatataa tcgatagatt gtgcacctg attgcccgc attatcgcga 6960
gccatttat acccatataa atcagcatcc atgttggaat ttaatcgcg cctcgagcaa 7020
gacgtttccc gttgaatatg gtcataaca ccccttgat tactgtttat gtaagcagac 7080
agttttattg ttcattgatg tatattttta tcttggtgca tgtaacatca gagattttga 7140
gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg ttattgtctc 7200
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 7260
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 7320
aaaaataggc gtatcacgag gccctttcgt c

```

7351

&lt;210&gt; 5

&lt;211&gt; 8008

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 5

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc accacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaa 540

```

tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgcc	ggcattatgc	ccagtacatg	accttatggg	actttccctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcgg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggatth	ccaagtctcc	accccatgta	780
cgtcaatggg	agtttgthtt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtatacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgthta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgtctgt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttacaatt	cacatataca	acaacgcgg	cccccggtgcc	1380
cgcagthttt	attaaacata	gcgtgggac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgthaacgg	ggagggcag	gtagtctgag	cagtactcgt	tgtgtccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgtcttaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgccaccgag	aagctgtggg	tgaccgtgta	ctacggcgtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcaa	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgacce	caacccccag	2160
gaggtggtgc	tggthaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaagg	gcagaaggag	tacgccttct	tctacaagct	ggacatcctc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtag	cttcgagccc	atcccccaacc	actactgcgc	ccccgcgggc	2580
ttcaccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	ccggccccctg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtgggtga	gcacccagct	gctgctgaac	2700
ggcagcctgg	ccgaggagga	ggtgggtgatc	cgcagcgtga	acttcaccga	caacgccaaag	2760
accatcatcg	tgcagctgaa	caccagcgtg	gagatcaact	gcacccgccc	caacaacaac	2820
acccgcaagc	gaattcgcat	ccagcgcggc	cccggccgcg	ccttcgtgac	catcggcaag	2880
atcggcaaca	tgcgccaggc	ccactgcaac	atcagccgcg	ccaagtggaa	caacaccctg	2940
aagcagatcg	acagcaagct	gcgcgagcag	ttcggcaaca	acaagaccat	catcttcaag	3000
cagagcagcg	gcggcgacce	cgagatcgtg	acccacagct	tcaactgcgg	cggcgagtct	3060
ttctactgca	acagcaccca	gctgttcaac	agcacctgg	tcaacagcac	ctggagcacc	3120
gagggcagca	acaacaccga	gggcagcgac	accatcacc	tgccctgccg	catcaagcag	3180
atcatcaaca	tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	3240
atccgctgca	gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	3300
aacgagagcg	agatcttccg	cctgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3360
ctgtacaagt	acaagggtgt	gaagatcgag	cccctgggcg	tggccccac	caaggccaag	3420
cgcgcgctgg	tgcagcgcga	gaagcgcgcc	gtgggcatcg	gcgcctgtt	cctgggcttc	3480
ctgggcgcgc	ccggcagcac	catgggcgcc	gccagcatga	ccctgaccgt	gcaggccccgc	3540
cagctgctga	gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggccag	3600
cagcacctgc	tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggccccg	catcctggcc	3660
gtggagcgct	acctgaagga	ccagcagctg	ctgggcatct	ggggctgcag	cggcaagctt	3720
atctgcacca	ccgcctgcc	ctggaacgcc	agctggagca	acaagagcct	ggagcagatc	3780
tggaaaccaca	ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgatc	3840

cacagcctga	tcgaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	3900
ctggacaagt	gggccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtacatc	3960
aagctgttca	tcatgatcgt	gggcggcctg	gtgggcctgc	gcatcgtgtt	cgccgtgctg	4020
agcatcgtga	accgcgtgcg	ccagggttac	agccccctga	gcttccagac	ccacctgccc	4080
atcccccgcg	gccccgaccg	ccccgagggc	atcgaggagg	agggcggcga	gcgcgaccgc	4140
gaccgcacga	tccgcctggt	gaacggcagc	ctggccctga	tctgggacga	cctgcgcagc	4200
ctgtgcctgt	tcagctacca	ccgcctgcgc	gacctgctgc	tgatcgtgac	ccgcacgtg	4260
gagctgctgg	gccgccgcgg	ctgggaggcc	ctgaagtact	ggtggaacct	gctgcagtac	4320
tggagccagg	agctgaagaa	cagcgccgtg	agcctgctga	acgccaccgc	catcatgggc	4380
ggcaagtggg	gcaagagcag	cgtgatcggc	tggcccgcgg	tgccgcgagcg	catgcgccgc	4440
gccgagcccc	ccgccgacgg	cgtgggcgcc	gtgagccgcg	acctggagaa	gcacggcgcc	4500
atcaccagca	gcaacaccgc	cgccaacaac	gccgcctgcg	cctggctgga	ggcccaggag	4560
gaggaggagg	tgggcttcgc	agtgaccgca	caggtggcac	tgccgcgcaat	gacctacaag	4620
gccgccgtgg	acctgagcca	cttcctgaag	gagaagggcg	gcctggaggg	cctgatccac	4680
agccagcgcc	gccaggacat	cctggacctg	tggatctacc	acaccagggg	ctacttcccc	4740
gactggcaga	actacacccc	cggccccggc	gtgcgctacc	ccctgacctt	cggctggtgc	4800
tacaagctgg	tgcccgtgga	gcccagacaag	gtggaggagg	ccaacaaggg	cgagaacacc	4860
agcctgctgc	accccgtgag	cctgcacggc	atggcagcac	ccgagcgcg	ggtgctggag	4920
tggcgcttcg	acagccgcct	ggccttccac	cacgtggccc	gcgagctgca	ccccgagtac	4980
ttcaagaact	gctgaggatc	cagatctgct	gtgccttcta	gltgccagcc	atctgttgtt	5040
tgccctccc	ccgtgccttc	cttgacctg	gaaggtgcc	ctcccactgt	cctttcctaa	5100
taaaatgagg	aaattgcac	gcattgtctg	agtaggtgtc	attctattct	ggggggtggg	5160
gtggggcagg	acagcaaggg	ggaggattgg	gaagacaata	gcaggcatgc	tggggatgcg	5220
gtgggctcta	tgggtaccca	ggtgctgaag	aattgacctg	gttcctcctg	ggccagaaag	5280
aagcaggcac	atcccttct	ctgtgacaca	ccctgtccac	gcccctggtt	cttagttcca	5340
gccccactca	taggacactc	atagctcagg	agggctccgc	cttcaatccc	acccgctaaa	5400
gtacttggag	cggtctctcc	ctccctcatc	agcccaccaa	accaaacct	gcctccaaga	5460
gtgggaagaa	attaaagcaa	gataggctat	taagtgcaga	gggagagaaa	atgcctccaa	5520
catgtgagga	agtaatgaga	gaaatcatag	aattttaagg	ccatcatggc	cttaatcttc	5580
cgcttctctg	ctcactgact	cgctgcgctc	ggctgcttcg	ctgcggcgag	cggtatcagc	5640
tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	5700
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	5760
ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agagggtggcg	5820
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	5880
tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	ttcttccctt	cgggaagcgt	5940
ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgctccaa	6000
gctgggctgt	gtgcacgaac	cccccgttca	gcccagaccg	tgcccttat	ccggtactaa	6060
tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	6120
caggattagc	agagcgagg	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	6180
ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	6240
cggaaaaaga	gltggtagct	cttgatccgg	caaacaacc	accgctggta	gcggtggttt	6300
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	6360
cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgttaaggga	ttttggtcat	6420
gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	gttttaaatc	6480
aatctaaagt	atatatgagt	aaacttgggtc	tgacagttac	caatgcttaa	tcagtgaggc	6540
acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcg	gggggggggg	6600
gcgctgaggt	ctgcctcgtg	aagaagggtg	tgctgactca	taccaggcct	gaatcgcccc	6660
atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	gctttgttgt	aggtggacca	6720
gttggtgatt	ttgaactttt	gctttgccac	ggaacgggtc	gcgttgtcgg	gaagatgcgt	6780
gatctgatcc	ttcaactcag	caaaagtctg	atttattcaa	caaagccgcc	gtccccgtcaa	6840
gtcagcgtaa	tgctctgcca	gtggttacaac	caattaacca	attctgatta	gaaaaactca	6900
tcgagcatca	aatgaaactg	caatttatcc	atatcaggat	tatcaatacc	atatttttga	6960
aaaagccgtt	tctgtaatga	aggagaaaac	tcaccgaggc	agttccatag	gatggcaaga	7020
tcctgggtatc	ggtctgcgat	tccgactcgt	ccaacatcaa	tacaacctat	taattttccc	7080
tcgtcaaaaa	taagggttatc	aagtgagaaa	tcaccatgag	tgacgactga	atccggtgag	7140



```

aatggcaaaa gcttatgcat ttctttccag acttggtcaa caggccagcc attacgctcg 7200
tcatcaaaat cactcgcatc aaccaaaccg ttattcattc gtgattgctc ctgagcgaga 7260
cgaaatacgc gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc 7320
aggaacactg ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc 7380
tggaatgctg ttttcccggg gatcgcatg gtgagtaacc atgcatcatc aggagtacgg 7440
ataaaatgct tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc 7500
tcatctgtaa catcattggc aacgctacct ttgccatggt tcagaaacaa ctctggcgca 7560
tcgggcttcc catacaatcg atagattgtc gcacctgatt gcccgacatt atcgcgagcc 7620
catttatacc catataaatc agcatccatg ttggaattta atcgcggcct cgagcaagac 7680
gtttcccggt gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt 7740
tttattgttc atgatgatat atttttatct tgtgcaatgt aacatcagag attttgagac 7800
acaacgtggc tttccccccc ccccattat tgaagcattt atcagggtta ttgtctcatg 7860
agcgataca tatttgaatg tatttagaaa aataaaciaa taggggttcc gcgcacattt 7920
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 7980
aataggcgta tcacgaggcc ctttcgtc 8008

```

&lt;210&gt; 6

&lt;211&gt; 6970

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 6

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgct gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacia 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtg cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgcc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```

gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga caccatgcgc 1920  
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040  
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 2100  
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 2160  
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220  
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280  
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340  
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400  
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520  
caggcctgcc ccaaggtgag ctctcgagccc atccccatcc actactgcgc ccccgccggc 2580  
ttcgccatcc tgaagtgcaa caacaagacc ttcaacggca ccggccctg caccaacgtg 2640  
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700  
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760  
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcaccgccc caacaacaac 2820  
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880  
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940  
aagcagatcg acagcaagct gcgcgagcag ttcggaaca acaagaccat catcttcaag 3000  
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgc cggcgagttc 3060  
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120  
gagggcagca acaacaccga gggcagcgac accatcacc tgccctgccg catcaagcag 3180  
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240  
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgc caacagcaac 3300  
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360  
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420  
cgccgcgtgg tgcagcgca gaagcgcgc gtgggcacg gcgcctgtt cctgggcttc 3480  
ctgggcgcgc ccggcagcac catgggcgcg gccagcatga ccctgaccgt gcaggcccgc 3540  
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600  
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660  
gtggagcgct acctgaagga ccagcagctg ctgggcacgt ggggctgcag cggcaagctt 3720  
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780  
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840  
cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900  
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtgagga 3960  
tccagatctg ctgtgccttc tagttgccag ccatctgttg ttgcccctc ccccgctgct 4020  
tccttgaccc tggaaggtgc cactcccact gtcccttctt aataaaatga ggaaattgca 4080  
tcgcattgtc tgagtaggtg tcattctatt ctggggggtg ggggtggggca ggacagcaag 4140  
ggggaggatt ggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgggtacc 4200  
caggtgctga agaattgacc cggttcctcc tgggccagaa agaagcaggc acatcccctt 4260  
ctctgtgaca caccctgtcc acgcccctgg ttcttagttc cagcccact cataggacac 4320  
tcatagctca ggagggtcc gccttcaatc ccaccgcta aagtacttgg agcggctctt 4380  
ccctcccctc tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4440  
aagataggct attaatgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4500  
gagaaatcat agaattttaa ggccatcatg gccttaatct tccgcttctt cgctcactga 4560  
ctcgtgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4620  
acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca 4680  
aaaggccagg aaccgtaaaa aggcgcgctt gctggcgctt tcccatagge tccgcccccc 4740  
tgacgagcat caaaaaatc gacgtcaag tcagaggtgg cgaaaccga caggactata 4800  
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 4860  
gcttaccgga tacctgtccg cctttctccc ttccgggaagc gtggcgcttt ctcatagctc 4920  
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 4980



```

accccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 5040
ggtaagacac gacttatcgc cactggcagc agccactggg aacaggatta gcagagcgag 5100
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 5160
aacagtatct ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 5220
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt ttttttggtt gcaagcagca 5280
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5340
cgctcagtgg aacgaaaact cacgttaagg gattttgggt atgagattat caaaaaggat 5400
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 5460
gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5520
tctatttcgt tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5580
tgaagaagg gttgctgact cataccaggc ctgaatcgcc ccacatcca gccagaaagt 5640
gaggagacca cgggtgatga gagctttggt gtaggtggac cagttgggtg ttttgaactt 5700
ttgctttgcc acggaacggg ctgcgttggt gggaagatgc gtgatctgat ccttcaactc 5760
agcaaaagtt cgatttatct aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc 5820
cagtggtaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5880
tgcaatttat tcatatcagg attatcaata ccataatctt gaaaaagccg tttctgtaat 5940
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 6000
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataagggtt 6060
tcaagtgaga aatcaccatg agtgacgact gaatccgggt agaatggcaa aagcttatgc 6120
atctctttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 6180
tcaaccaaac cgttatctat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 6240
ttaaaggagc aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6300
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6360
gggatcgtag tggtagtaaa ccatgcatca tcaggagtac ggataaaatg cttgatgggtc 6420
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6480
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6540
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6600
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6660
ctcataacac cccttgattt actgtttatg taagcagaca gttttattgt tcatgatgat 6720
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6780
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcgata catatttgaa 6840
tgtattttag aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6900
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6960
ccctttcgtc
6970

```

&lt;210&gt; 7

&lt;211&gt; 6787

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 7

```

tcgcgcggtt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatggtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540

```

tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtatacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatltatta	tttataaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcggttac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgcctcg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcag	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgccaccgag	aagctgtggg	tgaccgtgta	ctacggcggtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcca	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtggtgc	tgggtgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaaggt	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtgag	cttcgagccc	atccccaaac	actactgcgc	ccccgcccgc	2580
ttcaccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	ccggccccctg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcacccagct	gctgctgaac	2700
ggcagcctgg	ccgaggagga	ggtggtgatc	cgcagcgtga	acttcaccga	caacgccaag	2760
accatcatcg	tgcagctgaa	caccagcgtg	gagatcaact	gcacccgccc	caacaacaac	2820
acccgcaagc	gaattcgcat	ccagcgcggc	cccggccgcg	ccttcgtgac	catcggcaag	2880
atcggcaaca	tgcgccaggc	ccactgcaac	atcagccgcg	ccaagtggaa	caacaccctg	2940
aagcagatcg	acagcaagct	gcgcgagcag	ttcggcaaca	acaagaccat	catcttcaag	3000
cagagcagcg	gcggcgaccc	cgagatcgtg	accacagct	tcaactgcgg	cggcgagttc	3060
ttctactgca	acagcaccca	gctgttcaac	agcacctggt	tcaacagcac	ctggagcacc	3120
gagggcagca	acaacaccga	gggcagcgac	accatcaccc	tgccctgccg	catcaagcag	3180
atcatcaaca	tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	3240
atccgctgca	gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	3300
aacgagagcg	agatcttccg	cctgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3360
ctgtacaagt	acaaggtggt	gaagatcgag	cccctgggcg	tggccccac	caaggccaag	3420
cttaccgtgc	aggcccgcga	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	3480
cgcgccatcg	aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	3540
caggcccgc	tcctggccgt	ggagcgctac	ctgaaggacc	agcagctgga	gcagatctgg	3600
aaccacacca	cctggatgga	gtgggaccgc	gagatcaaca	actacaccag	cctgatccac	3660
agcctgatcg	aggagagcca	gaaccagcag	gagaagaacg	agcaggagct	gctggagctg	3720
gacaagtggg	ccagcctgtg	gaactggttc	aacatcacca	actggctgtg	gtgaggatcc	3780
agatctgctg	tgccttctag	ttgccagcca	tctgttgttt	gccccctccc	cgtgccttcc	3840

```

ttgaccctgg aaggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg 3900
cattgtctga gtaggtgtca ttctattctg gggggtggg tggggcagga cagcaagggg 3960
gaggattggg aagacaatag caggcatgct ggggatgcg tgggctctat gggtagccag 4020
gtgctgaaga attgacccgg ttctctctgg gccagaaaga agcaggcaca tccccttctc 4080
tgtgacacac cctgtccacg cccctgggtc ttagttccag cccactcat aggacactca 4140
tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttggagc ggtctctccc 4200
tccctcatca gccacccaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag 4260
ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 4320
aaatcataga attttaaggc catcatggcc ttaatcttcc gcttctctgc tcaactgactc 4380
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4440
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag gccagcaaaa 4500
ggccaggaac cgtaaaaaag ccgcgttgct ggcgtttttc cataggctcc gccccctga 4560
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4620
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4680
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4740
ctgtagggtat ctcagttcgg tgtagggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4800
ccccgttcag cccgaccgct gcgccttctc cggtaactat cgtcttgagt ccaacccggt 4860
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4920
tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4980
agtatttggt atctgcgctc tgctgaagcc agttacctc ggaaaaagag ttggtagctc 5040
ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat 5100
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc tttctacgg ggtctgacgc 5160
tcagtgaac gaaaactcac gttaagggat tttgggtcatg agattatcaa aaaggatctt 5220
cacctagatc cttttaaatt aaaaatgaag ttttaaatac atctaaagta tatatgagta 5280
aacttggctc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5340
atttcgttca tccatagtgt cctgactcgg gggggggggg cgctgaggtc tgcctcgtga 5400
agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5460
ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5520
ctttgccacg gaacggctctg cgttgctggg aagatgcgtg atctgacct tcaactcagc 5580
aaaagtctga ttattcaac aaagccgccc tcccgtcaag tcagcgtaat gctctgccag 5640
tgttacaacc aattaacca ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5700
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5760
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5820
ccgactcgtc caacatcaat acaacctatt aatttccct cgtcaaaaat aagggttatca 5880
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag cttatgcatt 5940
tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcacat 6000
accaaacctg tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgcgtgta 6060
aaaggacaat tacaacacag aatcgaatgc aaccggcgca ggaacactgc cagcgcacat 6120
acaatatatt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6180
atcgagtggt tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6240
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6300
acgctacctt tgccatgttt cagaaacaac tctggcgcac cgggcttccc atacaatcga 6360
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6420
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccggtt aatatggctc 6480
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6540
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tttccccccc 6600
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6660
atthagaaaa ataaacaaat aggggttccg cgcacatttc ccgaaaagt gccacctgac 6720
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6780
tttcgtc

```

6787

&lt;210&gt; 8

&lt;211&gt; 6607



&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 8

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttcccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tggcggttag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatgggggtc ccattttatta tttaaaatt cacaataaca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggtcgtcctg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccgtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800
cgccgccaca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatecgcg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcggtg 2040
cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaagggtgag cttcgagccc atccccaa accactgctg ccccgccggc 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggccctg caccaacgtg 2640
agcaccgtgc agtgaccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcaccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggaag 2880

```

atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940  
 aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000  
 cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060  
 ttctactgca acagcaccga gctgttcaac agcacctggg tcaacagcac ctggagcacc 3120  
 gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180  
 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240  
 atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300  
 aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360  
 ctgtacaagt acaagggtggg gaagatcgag cccctgggcg tggccccac caaggccaag 3420  
 cttaccgtgc agggccgcca gctgctgagc ggcacgtgc agcagcagaa caacctgctg 3480  
 cgcgccatcg agggccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 3540  
 cagggccgca tcctggcgtg ggagcgctac ctgaaggacc agcagctcga gtgaggatcc 3600  
 agatctgctg tgccctctag ttgccagcca tctgttgttt gcccctccc cgtgccttcc 3660  
 ttgacctgg aaggtgccac tcccactgtc ctttccctaat aaaatgagga aattgcatcg 3720  
 cattgtctga gtaggtgtca ttctattctg ggggggtggg tggggcagga cagcaagggg 3780  
 gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat ggggtaccag 3840  
 gtgctgaaga attgaccggt ttcctcctgg gccagaaaga agcaggcaca tccccttctc 3900  
 tgtgacacac cctgtccacg cccctgggtc ttagttccag cccactcat aggacactca 3960  
 tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttggagc ggtctctccc 4020  
 tccctcatca gccacacaaa ccaaaccctag cctccaagag tgggaagaaa ttaaagcaag 4080  
 ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 4140  
 aaatcataga attttaaggc catcatggcc ttaatcttcc gcttcctcgc tcaactgactc 4200  
 gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4260  
 gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag gccagcaaaa 4320  
 ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4380  
 cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag 4440  
 ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga cctgcccgt 4500  
 taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4560  
 ctgtaggtat ctcagttcgg tgtaggctgt tcgctccaag ctgggctgtg tgcacgaacc 4620  
 cccggttcag cccgaccgct gcgccttate cggtaactat cgtcttgagt ccaaccgggt 4680  
 aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4740  
 tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4800  
 agtatttggg atctgcgctc tgcgtgaagcc agttaccttc ggaaaaagag ttggtagctc 4860  
 ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca agcagcagat 4920  
 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4980  
 tcagtggaa gaaaactcac gttaagggat tttggatcat agattatcaa aaaggatctt 5040  
 cacctagatc cttttaaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5100  
 aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5160  
 atttcgttca tccatagttg cctgactcgg gggggggggg cgctgagggtc tgcctcgtga 5220  
 agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagttag 5280  
 ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5340  
 ctttgccacg gaacggctct cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5400  
 aaaagttcga tttattcaac aaagccgccg tcccgtaag tcagcgtaat gctctgccag 5460  
 tggtacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5520  
 aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5580  
 ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5640  
 ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aaggttatca 5700  
 agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag cttatgcatt 5760  
 tctttccaga cttgttcaac agggcagcca ttacgctcgt catcaaaatc actcgcattca 5820  
 accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5880  
 aaaggacaat tacaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5940  
 acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6000  
 atcgagtggt tgagtaacca tgcattcatca ggagtacgga taaaatgctt gatggtcgga 6060  
 agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6120  
 acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6180

tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6240  
 gcatccatgt tgggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6300  
 ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6360  
 tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tcccccccc 6420  
 cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6480  
 atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6540  
 gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6600  
 tttcgtc 6607

<210> 9

<211> 7042

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 9

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
 ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420  
 cccgcctggc tgaccgcccc acgacccccg ccattgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
 agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
 tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
 ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380  
 cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
 agcggctcat ggtcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga caccatgcgc 1920  
 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
 ggcattgctga tgatctgcag cgcaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040  
 cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgcca ggcctacgac 2100



accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160  
 gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220  
 cagatgcacg aggacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 2280  
 acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340  
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400  
 accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520  
 caggcctgcc ccaaggtgag ctctgagccc atccccaacc actactgcgc ccccgccggc 2580  
 ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccctg caccaacgtg 2640  
 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700  
 ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaa 2760  
 accatcatcg tgcagctgaa caccagcgtg gagatcaact gcaccgccc caacaacaac 2820  
 acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggaag 2880  
 atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940  
 aagcagatcg acagcaagct gcgcgagcag ttccggcaaca acaagaccat catcttcaag 3000  
 cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgc cggcgagtgc 3060  
 ttctactgca acagcaccca gctgttcaac agcacctggg tcaacagcac ctggagcacc 3120  
 gagggcagca acaacaccga gggcagcgac accatcacc tgccttgcgc catcaagcag 3180  
 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240  
 atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgc caacagcaac 3300  
 aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360  
 ctgtacaagt acaaggtggg gaagatcgag cccctgggcg tggcccccac caaggccaag 3420  
 cgccgcgtgg tgcagcgcga gaagcgcgc gtgggcatcg gcgcctgtt cctgggcttc 3480  
 ctgggcgcgc ccggcagcac catgggcgc gccagcatga ccctgaccgt gcaggcccgc 3540  
 cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600  
 cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660  
 gtggagcgt acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720  
 atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780  
 tggaaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840  
 cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900  
 ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960  
 aagctgttca tcatgatcgt gggcggcctg gtgggcctgc gcacgtgtt cgccgtgctg 4020  
 agcatctgag gatccagatc tgctgtgcct tctagtgtgc agccatctgt tgtttgcccc 4080  
 tccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaaat 4140  
 gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 4200  
 caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc 4260  
 tctatgggta cccaggtgct gaagaattga cccggttcct cctgggccag aaagaagcag 4320  
 gcacatcccc ttctctgtga cacacctgt ccacgcccct ggttcttagt tccagcccca 4380  
 ctcataggac actcatagct caggagggtc ccgccttcaa tcccaccgc taaagtactt 4440  
 ggagcggctc tccctccct catcagccca ccaaaccaaa cctagcctcc aagagtggga 4500  
 agaaattaaa gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg 4560  
 aggaagtaat gagagaaatc atagaatttt aaggccatca tggccttaat cttccgcttc 4620  
 ctgctcact gactcgctgc gctcggctgt tccgctgcgc cgagcgggat cagctcactc 4680  
 aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 4740  
 aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 4800  
 gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaacc 4860  
 gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 4920  
 tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 4980  
 ttctcatagc tcacgctgta ggtatctcag ttccggtgtag gtcgttcgct ccaagctggg 5040  
 ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgc ttatccggta actatcgtct 5100  
 tgagtccaac ccggttaagac acgacttatc gccactggca gcagccactg gtaacaggat 5160  
 tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgg ctaactacgg 5220  
 ctacactaga agaacagtat ttggtatctg cgtctgtctg aagccagtta ccttcggaaa 5280  
 aagagtgggt agctcttgat ccggcaaaaa aaccaccgct ggtagcgggt gtttttttgt 5340  
 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 5400

```

tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggatttttg tcatgagatt 5460
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 5520
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 5580
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctcggggggg gggggcgctg 5640
aggtctgcct cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc 5700
cagccagaaa gtgagggagc cacggttgat gagagctttg ttgtagggtg accagttggt 5760
gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg 5820
atccttcaac tcagcaaaaag ttcgatttat tcaacaaagc cgccgtcccc tcaagtcagc 5880
gtaatgctct gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc 5940
atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc 6000
cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg 6060
tateggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgtca 6120
aaaataaggt tatcaagtga gaaatcacca tgagtacga ctgaatccgg tgagaatggc 6180
aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacg ctcgatcatc 6240
aaatcactcg catcaaccaa accgttattc attcgtgatt ggcctgagc gagacgaaat 6300
acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac 6360
actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat 6420
gctgttttcc cggggatcgc agtggtagt aaccatgcat catcaggagt acggataaaa 6480
tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac catctcatct 6540
gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcacggggc 6600
ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg agccatttta 6660
taccatata aatcagcatc catgttgga tttaatcgcg gcctcgagca agacgtttcc 6720
cggtgaatat ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt 6780
gttcatgatg atatatTTTT atcttgtgca atgtaacatc agagattttg agaca'caacg 6840
tggttttccc cccccccca ttattgaagc atttatcagg gttattgtct catgagcgga 6900
tacatatattg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 6960
aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 7020
cgtatcacga ggccctttcg tc
7042

```

&lt;210&gt; 10

&lt;211&gt; 6859

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 10

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg bcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgTTTT ggcacaaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900

```



agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctggt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200  
 tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatgggggtc ccattttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380  
 cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500  
 agcggctcat ggctcgctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga caccatgcgc 1920  
 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
 ggcattgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040  
 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100  
 accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacce caacccccag 2160  
 gaggtgggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catgggtggag 2220  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280  
 accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacacca caaccaacagc 2340  
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400  
 accagcatcc gcggaagggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520  
 caggcctgcc ccaagggtgag cttcgagccc atcccccaacc actactgcgc ccccgccggc 2580  
 ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca cgggcccctg caccaacgtg 2640  
 agcaccgtgc agtgacacca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700  
 ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760  
 accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820  
 acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880  
 atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940  
 aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000  
 cagagcagcg gcggcgaccc cgagatcggt acccacagct tcaactgcgg cggcgagttc 3060  
 ttctactgca acagcacca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120  
 gagggcagca acaacaccga gggcagcgac accatcacc tgccctgccg catcaagcag 3180  
 atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240  
 atccgctgca gcagcaacat caccggcctg ctgctgacce gcgacggcgg caacagcaac 3300  
 aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360  
 ctgtacaagt acaagggtgt gaagatcgag cccctgggcg tggccccac caaggccaag 3420  
 cttaccgtgc agggccgcca gctgctgagc ggcacgtgac agcagcagaa caacctgctg 3480  
 cgcgccatcg agggccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 3540  
 caggcccgca tcctggccgt ggagcgctac ctgaaggacc agcagctgga gcagatctgg 3600  
 aaccacacca cctggatgga gtgggaccgc gagatcaaca actacaccag cctgatccac 3660  
 agcctgatcg aggagagcca gaaccagcag gagaagaacg agcaggagct gctggagctg 3720  
 gacaagtggg ccagcctgtg gaactggttc aacatcacca actggctgtg gtacatcaag 3780  
 ctgttcatca tgatcggtgg cggcctgggt ggcctgcgca tcgtgttcgc cgtgctgagc 3840  
 atctgaggat ccagatctgc tgtgccttct agttgccagc catctgttgt ttgccccctc 3900  
 cccgtgcctt ccttgacctt ggaagggtgc actcccactg tcctttccta ataaaatgag 3960  
 gaaattgcat cgcattgtct gagtaggtgt cattctatct tgggggggtg ggtggggcag 4020  
 gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 4080  
 atgggtaccc aggtgctgaa gaattgacct ggttctctct gggccagaaa gaagcaggca 4140  
 catccccttc tctgtgacac accctgtcca cgcccctggt tcttagttcc agccccactc 4200

```

ataggacact catagctcag gagggctccg ccttcaatcc caccgcgctaa agtacttgga 4260
gcggtctctc cctccctcat cagcccacca aaccaaacct agcctccaag agtgggaaga 4320
aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca acatgtgagg 4380
aagtaatgag agaaatcata gaattttaag gccatcatgg ccttaatctt ccgcttcctc 4440
gctcactgac tcgctgcgct cggctcgctc gctgcggcga gcggtatcag ctcactcaaa 4500
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4560
aggccagcaa aaggccagga accgtaaaaa ggccgcgctt ctggcgcttt tccataggct 4620
ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 4680
aggactataa agataccagg cgtttccccc tggaaagctc ctcgtgcgct ctctgttccc 4740
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc 4800
tcatagctca cgctgtaggc atctcagttc ggtgtaggct gttcgtcca agctgggctg 4860
tgtgcacgaa cccccgctt agcccgaccg ctgcgcctta tccggtaact atcgtcttga 4920
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4980
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta 5040
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 5100
agttggtagc tcttgatccg gcaaacaaac caccgctggg agcgggtggg tttttgtttg 5160
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 5220
ggggtctgac gctcagtga acgaaaactc acgttaaggg attttgggtc tgagattatc 5280
aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag 5340
tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc 5400
agcgatctgt ctatttcgtt catccatagt tgccgtactc gggggggggg ggcgctgagg 5460
tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag 5520
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttgggtgat 5580
tttgaacttt tgctttgcca cggaacggtc tgcgttgctg ggaagatgcg tgatctgac 5640
cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5700
atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc atcgagcatc 5760
aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5820
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5880
cggctctgca tcccgactcg tccaacatca atacaacct ttaatttccc ctcgtaaaaa 5940
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 6000
agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc gtcataaaaa 6060
tactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg 6120
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaaact 6180
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6240
gttttcccg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6300
ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgacct ctcatctgta 6360
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6420
ccatacaate gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6480
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccg 6540
tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag ttttattgtt 6600
catgatgata tatttttata ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6660
ctttccccc cccccatta ttgaagcatt tatcaggggt attgtctcat gagcggtac 6720
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa 6780
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6840
atcacgaggc cctttcgtc
6859

```

&lt;210&gt; 11

&lt;211&gt; 6965

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 11

```

tcgcgcggttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gtccatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg ttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatc tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga acaccatgcg 1920
cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgtggggca ccatgctgct 1980
gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040
gcccgtgtgg aaggaggcca ccaccacct gctctgcgc agcgacgcca aggcctacga 2100
caccgaggtg cacaacgtgt gggccacca cgcctgcgtg cccaccgacc ccaaccccc 2160
ggaggtggtg ctggtgaacg tgaccgagaa cttcgacatg tggaagaacg acatggtgga 2220
gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280
gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340
cagcagcggc cgcattgatca tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400
caccagcatc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460
ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520
ccaggcctgc cccaagggtg gcttcgagcc catccccaac cactactgcg ccccgcccg 2580
cttcaccatc ctgaagtgc aaggacaaga gttcaacggc aagggccct gcaccaacgt 2640
gagcaccgtg cagtgcacc acggcatccg ccccggtgtg agcaccagc tgctgctgaa 2700
cgaccgcctg gccgaggagg aggtggtgat ccgcagcgc aacttcgccc acaacgcaa 2760
ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcacccgcc ccaacaacaa 2820
caccgcaag agcatccaca tcggccccgg ccgcgccttc tacaccaccg gcgagatcat 2880
cggcgacatc cggcaggccc actgcaacct gagccgcgc aagtggaaac acaccctgaa 2940
caagatcgtg atcaagctgc gcgagcagtt cggcaacaag accatcgtgt tcaagcacag 3000
cagcggcggc gaccccgaga tcgtgacca cagcttcaac tgcggcgcg agttcttcta 3060
ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgagg 3120
cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180

```



caacatgtgg cagaaggtgg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240  
 ctgcagcagc aacatcacccg gcctgctgct gacccgcgac ggcggcaaca gcaacaacga 3300  
 gagcgagatc ttccgcctgg gcggcgggca catgcgcgac aactggcgca gcgagctgta 3360  
 caagtacaag gtggtgaaga tcgagcccc gggcgtggcc cccaccaagg ccaagcgccg 3420  
 cgtggtgcag cgcgagaagc gcgcgctggg catcggcgcc ctgttcctgg gcttcctggg 3480  
 cgccgcgggc agcaccatgg gcgcgcgcag catgaccctg accgtgcagg cccgccagct 3540  
 gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600  
 cctgctgcag ctgaccgtgt ggggcacaa gcagctgcag gcccgcaccc tggccgtgga 3660  
 gcgctacctg aaggaccagc agctgctggg catctggggc tgcagcggca agcttatctg 3720  
 caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780  
 ccacaccacc tggatggagt gggaccgcga gatcaacaac tacaccagcc tgatccacag 3840  
 cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900  
 caagtgggccc agcctgtgga actggttcaa catcaccaac tggctgtggt gaggatccag 3960  
 atctgctgtg ccttctagt ggcagccatc tgttgtttgc cctcccccg tgccttctt 4020  
 gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca 4080  
 ttgtctgagt aggtgtcatt ctattctggg ggggtggggtg gggcaggaca gcaaggggga 4140  
 ggattgggaa gacaatagca ggcattgctg ggatgcgggtg ggctctatgg gtaccaggt 4200  
 gctgaagaat tgacccggtt cctcctgggc cagaaagaag caggcacatc ccttctctg 4260  
 tgacacaccc tgtccacgcc cctggttctt agttccagcc ccactcatag gacactcata 4320  
 gctcaggagg gctccgcctt caatcccacc cgctaaagta cttggagcgg tctctccctc 4380  
 cctcatcagc ccaccaaacc aaacctagcc tccaagagtg ggaagaaatt aaagcaagat 4440  
 aggctattaa gtgcagaggg agagaaaatg cctccaacat gtgaggaagt aatgagagaa 4500  
 atcatagaat ttttaaggcca tcatggcctt aatcttcgc ttcctcgtc actgactcgc 4560  
 tgcgctcggc cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc 4620  
 tatccacaga atcaggggat aacgcaggaa agaactgtg agcaaaaggc cagcaaaagg 4680  
 ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 4740  
 agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4800  
 accaggcgtt tccccctgga agctccctcg tgcgtctcc tgttccgacc ctgccgctta 4860  
 ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4920  
 gtaggtatct cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 4980  
 ccgttcagcc cgaccgtgc gccttatccg gtaactatcg tcttgagtc aaccgggtaa 5040  
 gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 5100  
 taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag 5160  
 tattttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 5220  
 gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 5280  
 cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5340  
 agtggaaacga aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca 5400  
 cctagatcct tttaaattaa aaatgaagt ttaaataaat cttaaagtata tatgagtaaa 5460  
 cttggtctga cagttaccaaa tgcttaataa gtgaggcacc tatctcagcg atctgtctat 5520  
 ttcgttcatc catagttgcc tgactcgggg gggggggggc ctgagggtctg cctcgtgaag 5580  
 aagggtgttg tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5640  
 agccacgggt gatgagagct ttgtttagg tggaccagtt ggtgattttg aacttttgct 5700  
 ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa 5760  
 aagttcgatt tattcaacaa agccgcgcgc ccgtcaagtc agcgtaatgc tctgccagtg 5820  
 ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5880  
 tttattcata tcaggattat caataccata tttttgaaaa agcgtttct gtaatgaagg 5940  
 agaaaactca ccgaggcagt tccataggat ggcaagatcc tggatcgggt ctgcgattcc 6000  
 gactcgtcca acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag 6060  
 tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 6120  
 tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 6180  
 caaacggtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 6240  
 aggacaatta caaacaggaa tcgaatgcaa ccggcgagc aacactgcca gcgcatcaac 6300  
 aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6360  
 cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6420  
 aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6480

```

gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6540
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6600
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6660
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6720
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt cccccccccc 6780
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6840
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6900
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6960
tcgtc

```

&lt;210&gt; 12

&lt;211&gt; 6784

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 12

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tggggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcggtg 2040

```

```

cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcca ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggatgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacacca caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaagggtgag ctctcagccc atccccaacc actactgcgc ccccgccggc 2580
ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca agggccctg caccaacgtg 2640
agcaccgtgc agtgaccca cggcatccgc cccgtggtga gcaccagct gctggttacg 2700
ggtaacctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgcga caacgccaag 2760
gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcaccgccc caacaacaac 2820
acccgcaaga gcatccacat cggccccggc cgcgccttct acaccaccg cgagatcatc 2880
ggcgacatcc gccaggccca ctgcaacctg agccgcgcca agtggaacga caccctgaac 2940
aagatcgtga tcaagctgcg cgagcagttc ggcaacaaga ccatcgtgtt caagcacagc 3000
agcggcggcg accccgagat cgtgaccac agcttcaact gcggcggcga gttcttctac 3060
tgcaacagca cccagctgtt caacagcacc tggttcaaca gcacctggag caccgagggc 3120
agcaacaaca ccgagggcag cgacaccatc accctgccct gccgcatcaa gcagatcatc 3180
aacatgtggc agaagggtgg caaggccatg tacgcccccc ccatcagcgg ccagatccgc 3240
tgcagcagca acatcaccgg cctgctgctg acccgcgacg gcggcaacag caacaacgag 3300
agcgagatct tccgcctggg cggcgggcag atgcgcgaca actggcgag cgagctgtac 3360
aagtacaagg tggatgaagat cgagccctg ggctggtccc ccaccaaggc caagcttacc 3420
gtccaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 3480
atcgaggccc agcagcacct gctgcagctg accgtgtggg gcacacagca gctgcaggcc 3540
cgcacctg cgtggagcg ctacctgaag gaccagcagc tgctcgagca gatctggaac 3600
cacaccacct ggatggagtg ggaccgcgag atcaacaact acaccagcct gatccacagc 3660
ctgatcgagg agagccagaa ccagcacgag aagaacgagc aggagctgct ggagctggac 3720
aagtgggcca gcctgtggaa ctggttcaac atcaccaact ggctgtggtg aggatccaga 3780
tctgctgtgc cttctagtgt ccagccatct gttgtttgcc cctccccgt gccttccttg 3840
accctggaag gtgccactcc cactgtcctt tcttaataaa atgaggaaat tgcacgcat 3900
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 3960
gattgggaag acaatagcag gcacgtgtgg gatgcggtgg gctctatggg taccaggtg 4020
ctgaagaatt gaccgggttc ctccctgggc agaaagaagc aggcacatcc ctttctctgt 4080
gacacacct gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag 4140
ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggc ctctccctcc 4200
ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4260
ggctattaag tgcagagggg gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4320
tcatagaatt ttaaggccat catggcctta atcttccgct tctctgctca ctgactcgct 4380
gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg taatacgggt 4440
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 4500
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggcctccgc cccctgacga 4560
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 4620
ccaggcgttt cccctggaa gctccctcgt gcgctctct gttccgaccc tgccgcttac 4680
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 4740
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 4800
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag 4860
acacgactta tgcgacttg cagcagccac tggtaacagg attagcagag cgaggatagt 4920
aggcgggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt 4980
atlttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 5040
atccggcaaa caaacaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 5100
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 5160
gtggaacgaa aactcacgtt aagggtatct ggtcatgaga ttatcaaaaa ggatcttcac 5220
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 5280
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 5340

```



```

tcgttcatcc atagttgcct gactcggggg gggggggcgc tgaggtctgc ctctgaaga 5400
aggtgttgct gactcatacc aggcctgaat cgcccatca tccagccaga aagtgagga 5460
gccacgggtg atgagagctt tgttgtaggt ggaccagttg gtgattttga acttttgctt 5520
tgccacggaa cggctctgct tgtcgggaag atgcgtgatc tgatccttca actcagcaa 5580
agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtgt 5640
tacaaccaat taaccaatc tgattagaaa aactcatcga gcatcaaag aaactgcaat 5700
ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 5760
gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 5820
actcgtccaa catcaatata acctattaat tccccctcgt caaaaataag gttatcaagt 5880
gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 5940
ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcatacaac 6000
aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa 6060
ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag cgcatacaac 6120
atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc 6180
gcagtgggtg gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcggaaga 6240
ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 6300
ctacctttgc catgtttcag aaacaactct ggccatcgg gcttcccata caatcgatag 6360
attgtcgcac ctgattgccc gacattatcg cgagccatt tataccata taaatcagca 6420
tccatgttgg aatttaatcg cggcctcgag caagacgttt ccggttgaat atggctcata 6480
acacccttg tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt 6540
ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttt ccccccccc 6600
cattattgaa gcatttatca ggttattgt ctcatgagcg gatacatatt tgaatgtatt 6660
tagaaaaata acaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 6720
taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt 6780
cgtc
6784

```

&lt;210&gt; 13

&lt;211&gt; 7037

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 13

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaggt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt ttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080

```



tcttatgcat gctataactgt ttttggcttg gggcctatac acccccgcctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaate cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380  
 cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcggttac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctgggtc ccatgcctcc 1500  
 agcggctcat ggtcgcctcg cagctccttg ctccaaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat ggggtctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga acaccatgcg 1920  
 cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgtggggcga ccatgctgct 1980  
 gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040  
 gcccggtgtg aaggaggcca ccaccaccct gctctgcgcc agcgacgcca aggcctacga 2100  
 caccgaggtg cacaacgtgt gggccaccga cgctgcgtg cccaccgacc ccaaccccca 2160  
 ggaggtggtg ctggtgaacg tgaccgagaa ctctgacatg tggagaacg acatggtgga 2220  
 gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280  
 gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340  
 cagcagcggc cgcagatca tggagaagg cgagatcaag aactgcagct tcaacatcag 2400  
 caccagcacc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460  
 ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaacacca gcgtgatcac 2520  
 ccaggcctgc cccaagggtga gcttcgagcc catccccaac cactactgcg ccccgccgg 2580  
 ctccaccatc ctgaagtga aggacaagaa gttcaacggc aaggggccct gcaccaacgt 2640  
 gagcaccgtg cagtgcacc acggcatccg ccccggtgtg agcaccagc tgctgctgaa 2700  
 cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc aacttcgccc acaacgcca 2760  
 ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcaccgccc ccaacaaca 2820  
 caccgcaag agcatccaca tcggccccgg ccgcgccttc tacaccaccg gcgagatcat 2880  
 cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaaac acaccctgaa 2940  
 caagatcgtg atcaagctgc gcgagcagtt cggcaacaag accatcgtgt tcaagcacag 3000  
 cagcggcggc gaccccgaga tcgtgaccca cagcttcaac tgcggcgggc agttcttcta 3060  
 ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgagg 3120  
 cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180  
 caacatgtgg cagaagggtg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240  
 ctgcagcagc aacatcaccg gcctgctgct gaccgcgcac ggcggcaaca gcaacaacga 3300  
 gagcgagatc ttccgcctgg gcggcggcga catgcgcgac aactggcgca gcgagctgta 3360  
 caagtacaag gtggtgaaga tcgagccctt gggcggtggc cccaccaagg ccaagcgccg 3420  
 cgtggtgcag cgcgagaagc gcgccgtggg catcggcgcc ctgttccctg gcttccctgg 3480  
 cgccgcccgc agcaccatgg gcgccgccag catgaccctg accgtgcagg cccgccagct 3540  
 gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600  
 cctgctgcag ctgaccgtgt ggggcatcaa gcagctgcag gcccgcatcc tggccgtgga 3660  
 gcgctacctg aaggaccagc agctgctggg catctggggc tgcagcgga agcttatctg 3720  
 caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780  
 ccacaccacc tggatggagt gggaccgcga gatcaacaac tacaccagcc tgatccacag 3840  
 cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900  
 caagtgggccc agcctgtgga actggttcaa catcaccac tggctgtggt acatcaagct 3960  
 gttcatcatg atcgtggggc gcctggtggg cctgcgcac gtgttcgccg tgctgagcat 4020  
 ctgaggatcc agatctgctg tgccttctag ttgccagcca tctgttgttt gcccctcccc 4080  
 cgtgccttcc ttgaccctgg aaggtgccac tccactgtc ctttccctaat aaaatgagga 4140  
 aattgcatcg cattgtctga gtaggtgtca ttctattctg ggggggtggg tggggcagga 4200  
 cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat 4260  
 gggtagccag gtgctgaaga attgaccggg ttctctctgg gccagaaaga agcaggcaca 4320  
 tccccctctc tgtgacacac cctgtccacg cccctggttc ttagttccag cccactcat 4380

```

aggacactca tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttgaggc 4440
gggtctctccc tccctcatca gccacccaaa ccaaacctag cctccaagag tgggaagaaa 4500
ttaaagcaag ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa 4560
gtaatgagag aaatcataga attttaaggc catcatggcc ttaatcttcc gcttcctcgc 4620
tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4680
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 4740
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4800
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag 4860
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4920
ccctgccgct taccggatac ctgtccgcct ttctcccttc ggaagcgtg gcgctttctc 4980
atagctcacg ctgtagggtat ctgagttcgg tgtaggctgt tcgctccaag ctgggctgtg 5040
tgcacgaacc ccccgttcag cccgaccgct gcgccttate cggtaactat cgtcttgagt 5100
ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 5160
gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca 5220
ctagaagaac agtatattgt atctgcgctc tgcgtgaagc agttaccttc ggaaaaagag 5280
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca 5340
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5400
ggcttgacgc tcagtggaac gaaaactcac gttaagggat tttgggtcatg agattatcaa 5460
aaaggatctt cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta 5520
tatatgagta aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag 5580
cgatctgtct atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc 5640
tgcctcgtga agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc 5700
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttgggtgattt 5760
tgaacttttg ctttgccacg gaacgggtctg cgttgctcggg aagatgcgtg atctgatcct 5820
tcaactcagc aaaagttcga tttattcaac aaagccgccg tcccgtaag tcagcgtaat 5880
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5940
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 6000
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 6060
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttccctc cgtcaaaaat 6120
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag 6180
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 6240
actcgcatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 6300
atcgctgtta aaaggacaat tacaacagg aatcgaatgc aaccggcgca ggaacactgc 6360
cagcgcatca acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt 6420
tttcccgggg atcgcagtgg tgagtaacca tgcacatca ggagtacgga taaaatgctt 6480
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6540
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6600
atacaatcga tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc 6660
atataaatca gcatccatgt tggaaatttaa tcgcggcctc gagcaagacg tttcccgttg 6720
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6780
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6840
ttcccccccc cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6900
at ttgaaatgt at ttgaaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6960
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 7020
cacgaggccc tttegtc

```

7037

&lt;210&gt; 14

&lt;211&gt; 6856

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 14

```

tcgcgcgttt  cggatgatgac  ggtgaaaacc  tctgacacat  gcagctcccg  gagacggtca  60
cagcttgtct  gtaagcggat  gccgggagca  gacaagcccc  tcagggcgcg  tcagcgggtg  120
ttggcgggtg  tcggggctgg  cttaactatg  cggcatcaga  gcagattgta  ctgagagtgc  180
accatatgcg  gtgtgaaata  ccgcacagat  gcgtaaggag  aaaataccgc  atcagattgg  240
ctattggcca  ttgcatacgt  tgtatccata  tcataatatg  tacatttata  ttggctcatg  300
tccaacatta  ccgccatgtt  gacattgatt  attgactagt  tattaatagt  aatcaattac  360
ggggtcatta  gttcatagcc  catatatgga  gttccgcgtt  acataactta  cggtaaattg  420
cccgcctggc  tgaccgcccc  acgacccccg  ccatttgacg  tcaataatga  cgtatgttcc  480
catagtaacg  ccaataggga  ctttccattg  acgtcaatgg  gtggagtatt  tacggtaaac  540
tgcccacttg  gcagtacatc  aagtgtatca  tatgccaaat  acgcccccta  ttgacgtcaa  600
tgacggtaaa  tggcccgctt  ggcattatgc  ccagtacatg  accttatggg  actttcctac  660
ttggcagtac  atctacgtat  tagtcatcgc  tattaccatg  gtgatgcggg  tttggcagta  720
catcaatggg  cgtggatagc  ggtttgactc  acgggggattt  ccaagtctcc  accccattga  780
cgtcaatggg  agtttgtttt  ggcacaaaaa  tcaacgggac  tttccaaaat  gtcgtaacaa  840
ctccgcccc  ttgacgcaaa  tgggcggtag  gcgtgtacgg  tgggaggtct  atataagcag  900
agctcgttta  gtgaaccgtc  agatcgccctg  gagacgccat  ccacgctgtt  ttgacctcca  960
tagaagacac  cgggaccgat  ccagcctccg  cggccgggaa  cgggtgcattg  gaacgcggat  1020
tccccgtgcc  aagagtgcag  taagtaccgc  ctatagactc  tataggcaca  cccctttggc  1080
tcttatgcat  gctatactgt  ttttggttg  gggcctatac  acccccgtt  ccttatgcta  1140
taggtgatgg  tatagcttag  cctatagggtg  tgggttattg  accattattg  accactcccc  1200
tattggtgac  gatactttcc  attactaatc  cataacatgg  ctctttgcca  caactatctc  1260
tattggttat  atgccaatat  tctgtccttc  agagactgac  acggactctg  tatttttaca  1320
ggatggggtc  ccattttatta  tttacaaatt  cacatataca  acaacgcgt  ccccggtgcc  1380
cgcagttttt  attaaacata  gcgtgggatc  tccacgcgaa  tctcgggtac  gtgttccgga  1440
catgggctct  tctccggtag  cggcggagct  tccacatccg  agccctggtc  ccatgcctcc  1500
agcggctcat  ggtcgtcctg  cagctccttg  ctctaacag  tggaggccag  acttaggcac  1560
agcacaatgc  ccaccaccac  cagtgtgcctg  cacaaggccg  tggcggtagg  gtatgtgtct  1620
gaaaatgagc  gtggagattg  ggctcgcacg  gctgacgcag  atggaagact  taaggcagcg  1680
gcagaagaag  atgcaggcag  ctgagttgtt  gtattctgat  aagagtcaga  ggtaactccc  1740
gttgcggtgc  tgttaacggt  ggagggcagt  gtagtctgag  cagtactcgt  tgctgccgcg  1800
cgcgccacca  gacataatag  ctgacagact  aacagactgt  tcctttccat  ggtcttttc  1860
tgcagtcacc  gtcgtcgaca  cgtgtgatca  gatatcgccg  ccgctctaga  caccatgcgc  1920
gtgaaggaga  agtaccagca  cctgtggcgc  tggggctggc  gctggggcac  catgctgctg  1980
ggcatgctga  tgatctgcag  cgccaccgag  aagctgtggg  tgaccgtgta  ctacggcgtg  2040
cccgtgtgga  aggaggccac  caccacctg  ctctgcgcca  gcgacgcaa  ggcctacgac  2100
accgaggtgc  acaacgtgtg  ggccaccac  gcctgcgtgc  ccaccgacc  caacccccag  2160
gaggtggtgc  tgggtgaacgt  gaccgagaac  ttcgacatgt  ggaagaacga  catggtggag  2220
cagatgcacg  aggacatcat  cagcctgtgg  gaccagagcc  tgaagccctg  cgtgaagctg  2280
acccccctgt  gcgtgagcct  gaagtgcacc  gacctgaaga  acgacaccaa  caccaacagc  2340
agcagcggcc  gcatgatcat  ggagaagggc  gagatcaaga  actgcagctt  caacatcagc  2400
accagcatcc  gcggcaagg  gcagaaggag  tacgccttct  tctacaagct  ggacatcatc  2460
cccatcgaca  acgacaccac  cagctacagc  ctgaccagct  gcaacaccag  cgtgatcacc  2520
caggcctgcc  ccaagggtgag  cttecgagcc  atccccacc  actactgcgc  ccccgccggc  2580
ttcgccatcc  tgaagtgcaa  ggacaagaag  ttcaacggca  agggccctg  caccaacgtg  2640
agcaccgtgc  agtgacccca  cggcatccgc  cccgtggtga  gcaccagct  gctggttacg  2700
ggtaacctgg  ccgaggagga  ggtggtgatc  cgcagcgcta  acttcgccga  caacgccaa  2760
gtgatcatcg  tgcagctgaa  cgagagcgtg  gagatcaact  gcaccgccc  caacaacaac  2820
acccgcaaga  gcatccacat  cggccccggc  cgcgccttct  acaccaccgg  cgagatcatc  2880
ggcgacatcc  gccaggcccc  ctgcaacctg  agccgcgcca  agtggaaacga  caccctgaac  2940
aagatcgtga  tcaagctgcg  cgagcagttc  ggcaacaaga  ccatcgtgtt  caagcacagc  3000
agcggcgggc  accccgagat  cgtgaccac  agcttcaact  gcggcgccga  gttcttctac  3060
tgcaacagca  cccagctgtt  caacagcacc  tgggttcaaca  gcacctggag  caccgagggc  3120
agcaacaaca  ccgagggcag  cgacaccatc  accctgcctt  gccgcatcaa  gcagatcatc  3180

```



```

aacatgtggc agaaggtggg caaggccatg tacgcccccc ccatcagcgg ccagatccgc 3240
tgcagcagca acatcaccgg cctgctgctg acccgcgacg ggggcaacag caacaacgag 3300
agcgagatct tccgcctggg cggcgggcgc atgcgcgcaca actggcgagc cgagctgtac 3360
aagtacaagg tggatgaagat cgagccccctg ggcgtggccc ccaccaaggc caagcttacc 3420
gtccaggccc gccagctgct gagcggcacg gtgcagcagc agaacaacct gctgcgcgcc 3480
atcgaggccc agcagcacct gctgcagctg accgtgtggg gcatcaagca gctgcaggcc 3540
cgcaccctgg ccgtggagcg ctacctgaag gaccagcagc tgctcgagca gatctggaac 3600
cacaccacct ggatggagtg ggaccgagc atcaacaact acaccagcct gatccacagc 3660
ctgatcgagg agagccagaa ccagcacgag aagaacgagc aggagctgct ggagctggac 3720
aagtgggcca gcctgtggaa ctgggttcaac atcaccaact ggctgtggta catcaagctg 3780
ttcatcatga tcgtggggcg cctgggtgggc ctgcgcacgc tgttcgccgt gctgagcacc 3840
tgaggatcca gatctgctgt gccttctagt tgccagccat ctggtgtttg cccctcccc 3900
gtgccttctt tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa 3960
attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac 4020
agcaaggggg aggattggga agacaatagc aggcattgct gggatgcggg gggctctatg 4080
ggtaccaggg tgctgaagaa ttgaccgggt tcctcctggg ccagaaagaa gcaggcacat 4140
ccccttctct gtgacacacc ctgtccacgc ccctgggttct tagttccagc cccactcata 4200
ggacactcat agctcaggag ggctccgcct tcaatcccac ccgctaaagt acttgagcgc 4260
gtctctccct ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaagaaat 4320
taaagcaaga taggctatta agtgagagg gagagaaaat gcctccaaca tgtgaggaag 4380
taatgagaga aatcatagaa ttttaaggcc atcatggcct taatcttccg cttcctcgct 4440
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 4500
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4560
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4620
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4680
actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac 4740
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4800
tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt 4860
gcacgaaccc ccggttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4920
caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4980
agcgagggtat gtaggcgggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 5040
tagaagaaca gtatgttggt tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 5100
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgtttgcaa 5160
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cttttgatct tttctacggg 5220
gtctgacgct cagtggaaacg aaaactcacg ttaagggtat ttgggtcatga gattatcaaa 5280
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 5340
atatgagtaa acttgggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5400
gatctgtcta tttcgttcat ccatagttgc ctgactcggg gggggggggc gctgaggtct 5460
gcctcgtgaa gaagggtgtg ctgactcata ccaggcctga atcgccccat catccagcca 5520
gaaagtgagg gagccacggg tgatgagagc tttgtttagt gtggaccagt tgggtatgtt 5580
gaacttttgc tttgccacgg aacggctctg gttgtcggga agatgcgtga tctgatcctt 5640
caactcagca aaagttcgat ttattcaaca aagccgcgt cccgtcaagt cagcgtaatg 5700
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa 5760
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5820
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5880
tctgcgattc cgactcgtec aacatcaata caacctatta atttccccct gtcaaaaaata 5940
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 6000
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 6060
ctcgcatcaa ccaaaccggt attcatctgt gattgcgcct gagcgagacg aaatacgcga 6120
tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgagc gaacactgcc 6180
agcgcatcaa caatattttc acctgaatca ggatattctt ctaataacct gaatgctgtt 6240
ttcccgggga tcgcagtggg gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6300
atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6360
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca 6420
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataacca 6480

```

tataaatcag catccatggt ggaatttaac cgcgccctcg agcaagacgt ttcccgttga 6540  
 atatggctca taacaccctt tgtattactg tttatgtaag cagacagttt tattgttcat 6600  
 gatgatatat ttttatcttg tgcaatgtaa catcagagat tttagagacac aacgtggctt 6660  
 tccccccccc ccattatttg aagcatttat cagggttatt gtctcatgag cggatacata 6720  
 tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6780  
 ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6840  
 acgagccctt ttcgtc 6856

<210> 15

<211> 6505

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 15

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgag gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gtcccgctt acataactta cggtaaattg 420  
 cccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
 agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200  
 tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380  
 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500  
 agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tgttaacggg ggagggcaggt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga caccatgcgc 1920  
 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
 ggcattgctga tgatctgcag cgccaccgag gctagcagcc tgaagtgcac cgacctgaag 2040  
 aacgacacca acaccaacag cagcagcggc cgcattgatca tggagaaggg cgagatcaag 2100  
 aactgcagct tcaacatcag caccagcatc cgcggcaagg tgcagaagga gtacgccttc 2160

ttctacaagc tggacatcat ccccatcgac aacgacacca ccagctacag cctgaccagc 2220  
 tgcaacacca gcgtgatcac ccaggcctgc cccaagggtga gcttcgagcc catccccaac 2280  
 cactactgcg cccccgccgg cttcgccatc ctgaagtga aggacaaga gttcaacggc 2340  
 aagggccccct gcaccaacgt gagcacggtg cagtgcaccc acggcatccg ccccggtggtg 2400  
 agcaccacagc tgctgctgaa cggtagcctg gccgaggagg aggtggtgat ccgcagcgct 2460  
 aacttcgccc acaacgcca ggtgatcatc gtgcagctga acgagagcgt ggagatcaac 2520  
 tgcaccgcc ccaacaacaa cccccgcaag agcatccaca tcggcccccg ccgcgccttc 2580  
 tacaccaccg gcgagatcat cggcgacatc cgccaggccc actgcaacct gagccgcgcc 2640  
 aagtggaaag acaccctgaa caagatcgtg atcaagctgc gcgagcagtt cggcaacaag 2700  
 accatcggtg tcaagcacag cagcggcggc gaccccgaga tcgtgaccca cagcttcaac 2760  
 tgccggcgcg agttcttcta ctgcaacagc acccagctgt tcaacagcac ctggttcaac 2820  
 agcacctgga gcaccgaggg cagcaacaac accgagggca gcgacaccat caccctgccc 2880  
 tgccgcatca agcagatcat caacatgtgg cagaagggtg gcaaggccat gtacgcccc 2940  
 cccatcagcg gccagatccg ctgcagcagc aacatcaccc gcctgctgct gaccgcgcac 3000  
 ggcggcaaca gcaacaacga gagcgagatc ttccgtcccg gcggcgcgca catgcgcgac 3060  
 aactggcgca gcgagctgta caagtacaag gtggtgaaga tcgagccctt gggcggtggc 3120  
 cccaccaagg ccaagcttac cgtccaggcc cgccagctgc tgagcggcat cgtgcagcag 3180  
 cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 3240  
 ggcataaagc agctgcaggc ccgcaccctg gccgtggagc gctacctgaa ggaccagcag 3300  
 cagctcgagc agatctggaa ccacaccacc tggatggagt gggaccgcga gatcaacaac 3360  
 tacaccagcc tgatccacag cctgatcgag gagagccaga accagcagga gaagaacgag 3420  
 caggagctgc tggagctgga caagtgggccc agcctgtgga actggttcaa catcaccaac 3480  
 tggctgtggt gaggatccag atctgctgtg ccttctagtt gccagccatc tgttgtttgc 3540  
 cctcccccg tgcttctctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa 3600  
 aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtgggtg 3660  
 gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatactgg ggatgcggtg 3720  
 ggctctatgg gtacccagggt gctgaagaat tgacccggtt cctcctgggc cagaaagaag 3780  
 caggcacatc ccttctcttg tgacacaccc tgtccacgcc cctggttctt agttccagcc 3840  
 ccactcatag gacactcata gctcaggagg gctccgcctt caatcccacc cgctaaagta 3900  
 cttggagcgg tctctccctc cctcatcagc ccaccaaac aaacctagcc tccaagagtg 3960  
 ggaagaaatt aaagcaagat aggctattaa gtgcagaggg agagaaaatg cctccaacat 4020  
 gtgaggaagt aatgagagaa atcatagaat tttaaggcca tcatggcctt aatcttccgc 4080  
 ttctctgctc actgactcgc tgcgctcggc cgttcggctg cggcgagcgg tatcagctca 4140  
 ctcaaaggcg gtaatacgggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 4200  
 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 4260  
 taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 4320  
 cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4380  
 tgttccgacc ctgcccgtta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 4440  
 gctttctcat agctcacgct gtaggtatct cagttcgggtg taggtcgttc gctccaagct 4500  
 gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 4560  
 tcttgagtcc aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 4620  
 gattagcaga gcgagggtat taggcggtgc tacagagttc ttgaagtggg ggcctaacta 4680  
 cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4740  
 aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 4800  
 tgtttgcaag cagcagatta cgcgcaaaa aaaaggatct caagaagatc ctttgatctt 4860  
 ttctacgggg tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgag 4920  
 attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagt tttaaataat 4980  
 ctaaagtata tatgagtaaa cttggtctga cagttaccaaa tgcttaatca gtgaggcacc 5040  
 tatctcagcg atctgtctat ttcggttcac catagttgcc tgactcgggg gggggggggc 5100  
 ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa tcgccccatc 5160  
 atccagccag aaagtgaggg agccacggtt gatgagagct ttgttgtagg tggaccagtt 5220  
 ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 5280  
 ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgccgtc ccgtcaagtc 5340  
 agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg 5400  
 agcatcaaat gaaactgcaa ttatttcata tcaggattat caataccata tttttgaaaa 5460



```

agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5520
tggtatcggc ctgcgattcc gactcgccca acatcaatac aacctattaa tttcccctcg 5580
tcaaaaataa gggtatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 5640
ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5700
tcaaaatcac tcgcatcaac caaaccgcta ttcattcgtg attgcgcctg agcgagacga 5760
aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgagg 5820
aacactgcc a ggcgcataac aatattttca cctgaatcag gatattcttc taatacctgg 5880
aatgctgttt tcccggggat cgcagtgggt agtaaccatg catcatcagg agtacggata 5940
aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 6000
tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcatcg 6060
ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 6120
ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga gcaagacgtt 6180
tcccgttgaa tatggctcat aacacccctt gtattactgt ttatgtaagc agacagtttt 6240
attgttcatg atgatataat tttatcttgt gcaatgtaac atcagagatt ttgagacaca 6300
acgtggcttt ccccccccc ccattattga agcatttatc agggttattg tctcatgagc 6360
ggatacatat ttgaatgtat ttgaaaaaat aaacaaatag gggttccgcg cacatttccc 6420
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaaat 6480
aggcgtatca cgaggccctt tcgtc 6505

```

&lt;210&gt; 16

&lt;211&gt; 6505

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 16

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aacgaccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgcgtgc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgcctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatgggggtc ccatttatta tttacaaatt cacatataca acaacgcctg ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgaggct tccacatccg agccctgggt ccatgcctcc 1500

```



agcggctcat ggctcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgcctgcccg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga caccatgcgc 1920  
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040  
cccgtgtgga aggaggccac caccaccctg ctctgcgcca ggcagccaa ggcctacgac 2100  
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160  
gaggtggtgc tgggtgaacgt gaccgagaac ttgcacatgt ggaagaacga catggtggag 2220  
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280  
acccccctgt gcgtagacct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340  
agcagcgcc gcattgatcat ggagaaggcg gagatcaaga actgcagctt caacatcagc 2400  
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccgc tagcatcaac 2520  
tgcacccgcc ccaacaacaa caccgcgaag agcatccaca tcggcccccg ccgcgccttc 2580  
tacaccaccg gcgagatcat cggcgacatc cgccaggccc actgcaacct gagccgcgc 2640  
aagtggaaac acaccctgaa caagatcgtg atcaagctgc gcgagcagtt cggcaacaag 2700  
accatcgtgt tcaagcacag cagcggcgcc gaccccgaga tcgtgaccca cagcttcaac 2760  
tgcggcgccg agttcttcta ctgcaacagc acccagctgt tcaacagcac ctggttcaac 2820  
agcacctgga gcaccgaggc cagcaacaac accgaggcca gcgacaccat caccctgccc 2880  
tgccgcatca agcagatcat caacatgtgg cagaaggctg gcaaggccat gtacgcccc 2940  
cccatcagcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 3000  
ggcggaaca gcaacaacga gagcgagatc ttccgtccgg gcggcgccga catgcgcgac 3060  
aactggcgca gcgagctgta caagtacaag gtggtgaaga tcgagccctt gggcgtggcc 3120  
cccaccaagg ccaagcttac cgtccaggcc cgccagctgc tgagcgccat cgtgcagcag 3180  
cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 3240  
ggcatcaagc agctgcaggc ccgcaccctg gccgtggagc gctacctgaa ggaccagcag 3300  
cagctcgagc agatctggaa ccacaccacc tggatggagt gggaccgcca gatcaacaac 3360  
tacaccagcc tgatccacag cctgatcgag gagagccaga accagcagga gaagaacgag 3420  
caggagctgc tggagctgga caagtgggcc agcctgtgga actggttcaa catcaccaac 3480  
tggctgtggt gaggatccag atctgctgtg ccttctagtt gccagccatc tgttgtttgc 3540  
ccctcccccg tgccttctt gaccctggaa ggtgccactc ccactgtcct ttccataata 3600  
aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtggggtg 3660  
gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcattgctg ggatgcggtg 3720  
ggctctatgg gtacccaggt gctgaagaat tgacccggtt cctcctgggc cagaaagaag 3780  
caggcacatc cccttctctg tgacacacc tgtccacgcc cctggttctt agttccagcc 3840  
ccactcatag gacactcata gctcaggagg gctccgcctt caatcccacc cgctaaagta 3900  
cttgagcgcg tctctccctc cctcatcagc ccaccaaacc aaacctagcc tccaagagt 3960  
ggaagaaatt aaagcaagat aggtatttaa gtgcagaggg agagaaaatg cctccaacat 4020  
gtgaggaagt aatgagagaa atcatagaat tttaaggcca tcatggcctt aatcttccgc 4080  
ttcctcgctc actgactcgc tgcgctcggc cgttcggctg cggcgagcg tatcagctca 4140  
ctcaaaggcg gtaatacggc tatccacaga atcaggggat aacgcaggaa agaacatgtg 4200  
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgcctg cgtttttcca 4260  
taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa 4320  
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4380  
tggtccgacc ctgcccgtta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 4440  
gctttctcat agctcacgct gtaggtatct cagttcgggt taggtcgttc gctccaagct 4500  
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 4560  
tcttgagtcc aaccggtaa gacacgact atcgccactg gcagcagcca ctggtaacag 4620  
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta 4680  
cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4740  
aaaaagagtt ggtagctctt gatccggcaa acaaacacc gctggtagcg gtggtttttt 4800

```

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatcct 4860
ttctacgggg tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgag 4920
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaataaat 4980
ctaaagtata tatgagtaaa cttgggtctga cagttaccaaa tgcttaatca gtgaggcacc 5040
tatctcagcg atctgtctat ttcgttcac ctagttgcc tgactcgggg gggggggggcg 5100
ctgagggtctg cctcgtgaag aagggtgttg tgactcatac caggcctgaa tcgccccatc 5160
atccagccag aaagtgaagg agccacggtt gatgagagct ttgttgtagg tggaccagtt 5220
gggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 5280
ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgcgctc ccgtcaagtc 5340
agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg 5400
agcatcaaat gaaactgcaa ttatttcata tcaggattat caataccata tttttgaaaa 5460
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5520
tggtatcggg ctgacgattcc gactcgtcca acatcaatac aacctattaa tttccctcgc 5580
tcaaaaataa gggtatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 5640
ggcaaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5700
tcaaaatcac tcgcatcaac caaaccgta ttcatcgtg attgcgcctg agcgagacga 5760
aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgccagg 5820
aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 5880
aatgctgttt tcccggggat cgcagtgggt agtaaccatg catcatcagg agtacggata 5940
aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 6000
tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcatcg 6060
ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 6120
ttatacccat ataaatcagc atccatgttg gaatttaate gcggcctcga gcaagacggt 6180
tcccggtgaa tatggtcat aacaccctt gtattactgt ttatgtaagc agacagtttt 6240
attgttcatg atgatataat tttatcttgt gcaatgtaac atcagagatt ttgagacaca 6300
acgtggcttt ccccccccc ccattattga agcattttatc agggttattg tctcatgagc 6360
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6420
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaaat 6480
aggcgtatca cgaggccctt tcgtc 6505

```

&lt;210&gt; 17

&lt;211&gt; 6646

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 17

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatggtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```

ctccgccccca ttgacgcaaa tgggcggttag gcgtgtacgg tgggaggtct atataagcag 900  
 agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctggt ttgacctcca 960  
 tagaagacac cgggaccgat ccagccctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatggggtc ccattttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380  
 cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggtgct tccacatccg agccctgggc ccatgcctcc 1500  
 agcggctcat ggtcgctcgg cagctccttg ctccataacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga cccatgcgc 1920  
 gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980  
 ggcagctga tgatctgcag cggcaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040  
 cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100  
 accgaggtgc acaacgtgtg ggccaccac gctcgtgc ccaccgacc caacccccag 2160  
 gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220  
 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280  
 accccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340  
 agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400  
 accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
 cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520  
 caggcctgcc ccaaggtgag cttcgagccc atccccacc actactgcgc ccccgccggc 2580  
 ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca agggccctg caccaacgtg 2640  
 agcaccgtgc agtgacccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700  
 ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgcca caacgccaag 2760  
 gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact gcaccgccc caacaacaac 2820  
 acccgcaaga gcatccacat cggccccggc cgcgccttct acaccaccgg cgagatcatc 2880  
 ggcgacatcc gccaggccca ctgcaacctg gctagcttct actgcaacag caccagctg 2940  
 ttcaacagca cctggttcaa cagcacctgg agcaccgagg gcagcaacaa caccgagggc 3000  
 agcgacacca tcacctgcc ctgcccgcac aagcagatca tcaacatgtg gcagaagggtg 3060  
 ggcaaggcca tgtacgcccc ccccatcagc ggccagatcc gctgcagcag caacatcacc 3120  
 ggctgctgc tgaccgcga cggcggaac agcaacaacg agagcgagat ctcccgctccg 3180  
 ggcgggcgcg acatgcgcga caactggcgc agcgagctgt acaagtacaa ggtggtgaag 3240  
 atcgagcccc tgggcgtggc cccaccaag gccaaagctta ccgccaggc cggccagctg 3300  
 ctgagcggca tctgtcagca gcagaacaac ctgctgcgcg ccacgaggc ccagcagcac 3360  
 ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg cccgcaccct ggccgtggag 3420  
 cgctacctga aggaccagca gcagctcgag cagatctgga accacaccac ctggatggag 3480  
 tgggaccgcg agatcaacaa ctacaccagc ctgatccaca gcctgatcga ggagagccag 3540  
 aaccagcagg agaagaacga gcaggagctg ctggagctgg acaagtgggc cagcctgtgg 3600  
 aactggttca acatcaccaa ctggctgtgg tgaggatcca gatctgctgt gccttctagt 3660  
 tgccagccat ctgttgtttg cccctcccc gtgccttctt tgaccctgga aggtgccact 3720  
 cccactgtcc tttcctaata aaatgaggaa attgcacgc attgtctgag taggtgtcat 3780  
 tctattctgg ggggtgggtt ggggcaggac agcaaggggg aggtattggga agacaatagc 3840  
 aggcagctg gggatgcggt gggctctatg ggtaccaggg tgctgaagaa ttgaccgggt 3900  
 tccctctggg ccagaaagaa gcaggcacat ccccttctct gtgacacacc ctgtccacgc 3960  
 cccctgggtct tagttccagc cccactcata ggacactcat agctcaggag ggctccgct 4020  
 tcaatccac ccgctaaagt acttgagcg gtctctccct cccatcatcag cccaccaaac 4080  
 caaacctagc ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg 4140



```

gagagaaaat gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc 4200
atcatggcct taatcttccg cttectcgct cactgactcg ctgcgctcgg tcgttcggct 4260
gcggcgagcg gtatcagctc actcaaaggg ggtaatacgg ttatccacag aatcagggga 4320
taacgcagga aagaacatgt gagcaaaaagg ccagcaaaaag gccaggaacc gtaaaaaggc 4380
cgcggttgcg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg 4440
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt tccccctgg 4500
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 4560
tctcccttcg ggaagcgtgg cgcttttctca tagctcacgc tgtagggtatc tcagttcggg 4620
gtaggtcggt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 4680
cgcttatcc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact 4740
ggcagcagcc actggtaaca ggattagcag agcgagggtat gtaggcgggtg ctacagagtt 4800
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct 4860
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 4920
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 4980
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg 5040
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 5100
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 5160
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5220
ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaagggtgtg ctgactcata 5280
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggg tgatgagagc 5340
tttggtgtag gtggaccagt tgggtgatttt gaacttttgc tttgccacgg aacggctctg 5400
gttgctcggg agatgcgtga tctgatcctt caactcagca aaagtctgat ttattcaaca 5460
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5520
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5580
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5640
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5700
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5760
acgactgaat ccggtgagaa tggcaaaaagc ttatgcattt ctttccagac ttgttcaaca 5820
ggccagccat tacgctcgtc atcaaaaatca ctgcgcatcaa ccaaaccggt attcattcgt 5880
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5940
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 6000
ggatattctt ctaatacctg gaatgctgtt ttcccgggga tcgcagtggg gagtaacct 6060
gcatcatcag gtagcggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6120
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6180
agaaacaact ctggcgcacg gggcttccca tacaatcgat agattgtcgc acctgattgc 6240
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6300
cgcggcctcg agcaagacgt ttcccggtga atatggctca taacaccctt tgtattactg 6360
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6420
catcagagat tttgagacac aacgtggctt tccccccccc cccattattg aagcatttat 6480
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6540
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6600
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6646

```

&lt;210&gt; 18

&lt;211&gt; 6670

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 18

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60

cagcttgtct	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgttaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaattg	420
cccgctggc	tgaccgcccc	acgacccccg	cccatgtacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggccccgct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgcg	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggatbt	ccaagtctcc	accccatgta	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgctcg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttacaaatt	cacatatata	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatacgcg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgccaccgag	aagctgtggg	tgaccgtgta	ctacggcggtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgccaa	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtggtgc	tggatgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaaggt	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccctcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtgag	cttcgagccc	atcccccaacc	actactgcgc	ccccgccggc	2580
ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	agggcccttg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	acttcgccga	caacgccaaag	2760
gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	gcaccgcgcc	caacaacaac	2820
accgcgaaga	gcatccacat	cggccccggc	cgcgccttct	acaccaccgg	cgagatcatc	2880
ggcgacatcc	gccaggccca	ctgcaacctg	agccgcgcca	agtggaaacga	caccctgaac	2940
aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	ccatcgtggt	caagcacagc	3000
agcggcggtg	accccgagat	cgtgaccac	agcttcaact	gcggcggtga	gttcttctac	3060
tgcaacagca	cccagctggt	caacagcacc	tggttcaaca	gcacctggag	caccgagggc	3120
agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	gccgcategc	tagcagcaac	3180
aacgagagcg	agatcttccg	tccgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3240
ctgtacaagt	acaagggtgt	gaagatcgag	cccctgggcg	tggccccac	caaggccaag	3300
cttaccgtcc	aggcccgcca	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	3360

```

cgcgccatcg aggccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 3420
caggcccgcg ccttgccgt ggagcgctac ctgaaggacc agcagcagct cgagcagatc 3480
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3540
cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3600
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtgagga 3660
tccagatctg ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgctgct 3720
tccttgaccc tggaagggtgc cactcccact gtcctttcct aataaaatga ggaaattgca 3780
tcgcattgtc tgagttaggtg tcattctatt ctgggggggtg ggggtggggca ggacagcaag 3840
ggggaggatt gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgggtacc 3900
cagggtgctga agaattgacc cggttcctcc tgggccagaa agaagcaggc acatcccctt 3960
ctctgtgaca caccctgtcc acgcccctgg ttcttagttc cagccccact cataggacac 4020
tcatagctca ggagggtccc gccttcaatc ccaccgcta aagtacttgg agcggtctct 4080
ccctccctca tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4140
aagataggct attaatgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4200
gagaaatcat agaattttaa ggccatcatg gccttaatct tccgcttctc cgctcactga 4260
ctcgtgcgc tcggtcgctt ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4320
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 4380
aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccatagge tccgcccccc 4440
tgacgagcat cacaaaaatc gacgctcaag tcagagggtg cgaaaccgca caggactata 4500
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 4560
gcttaccgga tacctgtccg cctttctccc ttccgggaagc gtggcgcttt ctcatagctc 4620
acgctgtagg tatctcagtt cgggtgtaggt cgctcgtccc aagctgggct gtgtgcacga 4680
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 4740
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4800
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 4860
aacagtatct ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttggtag 4920
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgg ttttttgttt gcaagcagca 4980
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5040
cgctcagtg aacgaaaact cacgttaagg gattttggct atgagattat caaaaaggat 5100
cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga 5160
gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5220
tctatctcgt tcatccatag ttgctgact cggggggggg gggcgctgag gtctgcctcg 5280
tgaagaaggt gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5340
gaggagacca cggttgatga gagctttggt gtaggtggac cagttggtga ttttgaactt 5400
ttgctttgcc acggaacggt ctgcgttgct gggaagatgc gtgatctgat ccttcaactc 5460
agcaaaagtt cgatttatct aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc 5520
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5580
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5640
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggctctgcg 5700
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataagggtta 5760
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaattggcaa aagcttatgc 5820
atctctttcc agacttgctt aacaggccag ccattacgct cgtcatcaaa atcactcgca 5880
tcaaccaaac cgttatctat tctgtattgc gcctgagcga gacgaaatac gcgatcgctg 5940
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6000
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccc 6060
gggatcgagc tgggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggct 6120
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6180
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6240
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6300
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccc ttgaatatgg 6360
ctcataacac cccttgatct actgtttatg taagcagaca gttttattgt tcatgatgat 6420
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6480
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6540
tgtatttaga aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6600
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6660

```

ccctttcgtc

6670

<210> 19

<211> 6709

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 19

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacaataaca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcggtg 2040
ccgctgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcca ggccctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacacca caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520

```



caggcctgcc	ccaaggtgag	cttcgagccc	atcccccaacc	actactgcgc	ccccgccggc	2580
ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	agggcccttg	caccaacgtg	2640
agcaccgtgc	agtgcacca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	acttcgccga	caacgccaaag	2760
gtgatcatcg	tgcagctgaa	cgagagcggtg	gagatcaact	gcaccgcgcc	caacaacaac	2820
acccgcaaga	gcatccacat	cggccccggc	cgcgccttct	acaccaccgg	cgagatcatc	2880
ggcgacatcc	gccaggccca	ctgcaacctg	agccgcgcca	agtggaaacga	cacctgaac	2940
aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	ccatcgtggt	caagcacagc	3000
agcggcgggc	accccgagat	cgtgaccac	agcttcaact	gcggcgggcg	gttcttctac	3060
tgcaacagca	cccagctggt	caacagcacc	tggttcaaca	gcacctggag	caccgagggc	3120
agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	gccgcatcaa	gcagatcatc	3180
aacatgtggc	agaaggtggg	caaggccatg	tacgcccccc	ccatcagcgg	ccagatccgc	3240
tgcagcagca	acatcacccg	cctgctgctg	acccgcgacg	gcggcaacag	caacaacgag	3300
agcgagatct	tccgtccggg	cggcggcgct	agcgccaagc	ttaccgtcca	ggcccgccag	3360
ctgctgagcg	gcatcgtgca	gcagcagaac	aacctgctgc	gcgccatcga	ggcccagcag	3420
cacctgctgc	agctgaccgt	gtggggcatc	aagcagctgc	aggcccgcac	cctggccgtg	3480
gagcgctacc	tgaaggacca	gcagcagctc	gagcagatct	ggaaccacac	cacctggatg	3540
gagtgggacc	gcgagatcaa	caactacacc	agcctgatcc	acagcctgat	cgaggagagc	3600
cagaaccagc	aggagaagaa	cgagcaggag	ctgctggagc	tggacaagtg	ggccagcctg	3660
tggaaactgg	tcaacatcac	caactggctg	tggtaggat	ccagatctgc	tgtgccttct	3720
agttgccagc	catctgttgt	ttgccccctc	cccgtgcctt	ccttgaccct	ggaaggtgcc	3780
actcccactg	tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	3840
cattctatct	tgggggggtg	ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	3900
agcaggcatg	ctggggatgc	ggtgggctct	atgggtaccc	aggtgctgaa	gaattgacc	3960
ggttcctcct	gggccagaaa	gaagcaggca	catccccctc	tctgtgacac	accctgtcca	4020
cgccccctgg	tcttagttcc	agccccactc	ataggacact	catagctcag	gagggctccg	4080
ccttcaatcc	caccgcgtaa	agtacttgga	gcggtctctc	cctccctcat	cagcccacca	4140
aaccaaacc	agcctccaag	agtgggaaga	aattaaagca	agataggcta	ttaagtgcag	4200
agggagagaa	aatgcctcca	acatgtgagg	aagtaatgag	agaaatcata	gaattttaag	4260
gccatcatgg	ccttaatctt	ccgcttctct	gctcactgac	tcgctgcgct	cggtcgttcg	4320
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggttaata	cggttatcca	cagaatcagg	4380
ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	4440
ggccgcggtg	ctggcggttt	tccataggct	ccgccccctc	gacgagcatc	acaaaaatcg	4500
acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	4560
tggaaagctc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	4620
ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	cgctgtaggt	atctcagttc	4680
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	4740
ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	4800
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	4860
gttcttgaag	tgggtggccta	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	4920
tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaac	4980
caccgctgg	agcgggtgggt	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	5040
atctcaagaa	gaccccttga	tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	5100
acgttaagg	attttggtea	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	5160
ttaaaaaatga	agtttttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	5220
ccaatgctta	atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	5280
tgccctgactc	gggggggggg	ggcgctgagg	tctgcctcgt	gaagaagggt	ttgctgactc	5340
ataccaggcc	tgaatcgccc	catcatccag	ccagaaagtg	agggagccac	ggttgatgag	5400
agctttgttg	taggtggacc	agttgggtgat	tttgaacttt	tgcttttgcca	cggaacggtc	5460
tgcggtgtcg	ggaagatgcg	tgatctgatc	cttcaactca	gcaaaagttc	gatttattca	5520
acaaagccgc	cgtcccgtca	agtcagcgta	atgctctgcc	agtgttacaa	ccaattaacc	5580
aattctgatt	agaaaaactc	atcgagcatc	aaatgaaact	gcaatttatt	catatcagga	5640
ttatcaatac	catatttttg	aaaaagccgt	ttctgtaatg	aaggagaaaa	ctcaccgagg	5700
cagttccata	ggatggcaag	atcctgggtat	cgggtctgcga	ttccgactcg	tccaacatca	5760
atacaaccta	ttaatttccc	ctcgtcaaaa	ataaggttat	caagtgagaa	atcaccatga	5820

```

gtgacgactg aatccggtga gaatggcaaa agcttatgca tttctttcca gacttggtca 5880
acaggccagc cattacgctc gtcatacaaa tcaactcgcat caaccaaacc gttattcatt 5940
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 6000
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6060
tcaggatatt cttctaatac ctggaatgct gttttcccg ggatcgcagt ggtgagtaac 6120
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgctc 6180
agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 6240
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6300
tgcccagacat tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt 6360
aatcgcggcc tcgagcaaga cgtttcccg tgaatatggc tcataacacc ccttgattta 6420
ctgtttatgt aagcagacag ttttattgtt catgatgata tattttttatc ttgtgcaatg 6480
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 6540
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaaca 6600
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6660
atcatgacat taacctataa aaataggcgt atcacgagc cctttcgtc 6709

```

&lt;210&gt; 20

&lt;211&gt; 6736

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 20

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aagacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```

```

gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccacgcct gcggtgccac cgaccccaac cccagggagg 2160
tgggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280
ccctgtgctg gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaagggt cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 2460
aggcctgccc caaggtgagc ttcgagccca tccccaacca ctactgcgcc cccgccggct 2520
tcgccatcct gaagtgcagg gacaagaagt tcaacggcaa gggcccttgc accaacgtga 2580
gcaccgtgca gtgcaccac ggcatccgcc ccgtggtgag caccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc aacaacaaca 2760
cccgcaagag catccacatc gggcccgggc gcgccttcta caccaccggc gagatcatcg 2820
gcgacatccg ccaggccccc tgcaacctga gccgcgccc aa gtggaacgac accctgaaca 2880
agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc aagcacagca 2940
gcggcgggcg ccccgagatc gtgaccacaa gcttcaactg cggcgggcag ttcttctact 3000
gcaacagcac ccagctgttc aacagcacct ggttcaacag caccctggagc accgagggca 3060
gcaacaacac cgaggggcagc gacaccatca ccctgccctg ccgcatcaag cagatcatca 3120
acatgtggca gaaggtgggc aaggccatgt acgcccccc catcagcggc cagatccgct 3180
gcagcagcaa catcaccggc ctgctgctga ccgcgacgg cggcaacagc aacaacgaga 3240
gcgagatctt ccgtccgggc ggcgcgaca tgcgcgacaa ctggcgagc gagctgtaca 3300
agtacaaggt ggtgaagatc gagcccttgg gcgtggcccc caccaaggcc aagcttaccg 3360
tccaggcccc ccagctgctg agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca 3420
tcgaggcccc gcagcacctg ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc 3480
gcaccctggc cgtggagcgc tacctgaagg accagcagca gctcgagcag atctggaacc 3540
acaccacctg gatggagtgg gaccgcgaga tcaacaacta caccagcctg atccacagcc 3600
tgatcgagga gagccagaac cagcaggaga agaacgagca ggagctgctg gagctggaca 3660
agtgggcccag cctgtggaac tggttcaaca tcaccaactg gctgtggtga ggatccagat 3720
ctgctgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccctg ccttccttga 3780
ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcategcatt 3840
gtctgagtag gtgtcattct attctggggg gtgggggtgg gcaggacagc aagggggagg 3900
attgggaaga caatagcagg catgctgggg atgcggtggg ctctatgggt acccaggtgc 3960
tgaagaattg acccggttcc tcctgggcca gaaagaagca ggcacatccc cttctctgtg 4020
acacaccctg tccacgcccc tggttcttag ttccagcccc actcatagga cactcatagc 4080
tcaggagggg tccgccttca atcccaccgg cttaaagtact tggagcggtc tctccctccc 4140
tcatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4200
gctattaagt gcagaggggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4260
catagaattt taaggccatg atttaaggcc atcatggcct taatcttccg cttcctcgct 4320
cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc 4380
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4440
ccagcaaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4500
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4560
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 4620
cctgcccgtt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgcttttctc 4680
tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt cgctccaagc tgggctgtgt 4740
gcacgaaccc ccggttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4800
caaccgggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4860
agcgagggat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact acggctacac 4920
tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 4980

```



```

tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 5040
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 5100
gtctgacgct cagtggaaacg aaaactcacg ttaagggatt ttggatcatga gattatcaaa 5160
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaataca tctaaagtat 5220
atatgagtaa acttggctctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5280
gatctgtcta ttctgttcat ccatagttgc ctgactcggg gggggggggc gctgaggtct 5340
gcctcgtgaa gaagggtgtg ctgactcata ccaggcctga atcgcccat catccagcca 5400
gaaagtgagg gagccacggg tgatgagagc ttgtttgtag gtggaccagt tgggtgatttt 5460
gaacttttgc ttgcccacgg aacggctctg gttgtcggga agatgcgtga tctgatcctt 5520
caactcagca aaagtctgat ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg 5580
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa 5640
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5700
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5760
tctgcgattc cgactcgtcc aacatcaata caacctatta atttccctc gtcaaaaata 5820
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5880
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5940
ctcgcaccaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 6000
tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgag gaacactgcc 6060
agcgcaccaa caatattttc acctgaatca ggatattctt ctaataacctg gaatgctgtt 6120
ttcccgggga tcgcagtggg gagtaaccat gcatcatcag gactacggat aaaatgcttg 6180
atggctcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6240
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca 6300
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca 6360
tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga 6420
atatggctca taacaccctt tgtattactg tttatgtaag cagacagttt tattgttcat 6480
gatgatatat ttttatcttg tgcaatgtaa catcagagat tttgagacac aacgtggctt 6540
tcccccccc cccattattg aagcatttat cagggttatt gtctcatgag cggatacata 6600
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6660
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6720
acgaggccct ttcgtc 6736

```

&lt;210&gt; 21

&lt;211&gt; 6697

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 21

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780

```

cgtcaatggg	agtttgTTTT	ggcaccAAAA	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtggggtatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcag	gtagtctgag	cagtactcgt	tgctgcccg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcctttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgctggg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	accacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcat	gatcatggag	aaggcgagga	tcaagaactg	cagcttcgct	agcaccagct	2400
gcaacaccag	cgtgatcacc	caggcctgcc	ccaagggtgag	cttcgagccc	atcccccaacc	2460
actactgcgc	ccccgcgggc	tctgccatcc	tgaagtgcga	ggacaagaag	ttcaacggca	2520
agggccccctg	caccaacgtg	agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	2580
gcacccagct	gctgctgaac	ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	2640
acttcgccga	caacgccaa	gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	2700
gcacccgccc	caacaacaac	accgcgaaga	gcattccacat	cggccccggc	cgcgccttct	2760
acaccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacctg	agccgcgcca	2820
agtggaaacga	caccctgaac	aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	2880
ccatcgtgtt	caagcacagc	agcggcgggc	accccgagat	cgtgaccac	agcttcaact	2940
gcggcgggcga	gttctttctac	tgcacacagca	cccagctgtt	caacagcacc	tgggttcaaca	3000
gcacctggag	caccgagggc	agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	3060
gccgcacaa	gcagatcatc	aacatgtggc	agaagggtggg	caaggccatg	tacgcccccc	3120
ccatcagcgg	ccagatccgc	tgcagcagca	acatcaccgg	cctgctgctg	acccgcgacg	3180
gcggcaacag	caacaacgag	agcgagatct	tccgtccggg	cggcgggcag	atgcgcgaca	3240
actggcgag	cgagctgtac	aagtacaagg	tgggtgaagat	cgagcccttg	ggcgtggccc	3300
ccaccaaggc	caagcttacc	gtccaggccc	gccagctgct	gagcggcatc	gtgcagcagc	3360
agaacaacct	gctgcgcgcc	atcgaggccc	agcagcacct	gctgcagctg	accgtgtggg	3420
gcataagca	gctgcaggcc	cgcaccctgg	ccgtggagcg	ctacctgaag	gaccagcagc	3480
agctcgagca	gatctgggaa	cacaccacct	ggatggagtg	ggaccgcgag	atcaacaact	3540
acaccagcct	gatccacagc	ctgatcgagg	agagccagaa	ccagcaggag	aagaacgagc	3600
aggagctgct	ggagctggac	aagtgggcca	gcctgtggaa	ctgggttcaac	atcaccaact	3660
ggctgtgggtg	aggatccaga	tctgctgtgc	cttctagtgtg	ccagccatct	gttgtttgcc	3720
cctcccccg	gccttccttg	accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	3780
atgaggaaat	tgcacgcgat	tgtctgagta	ggtgtcatte	tattctgggg	ggtgggggtg	3840
ggcaggacag	caagggggag	gattggggaag	acaatagcag	gcattgctggg	gatgcgggtg	3900
gctctatggg	taccaggtg	ctgaagaatt	gacccgggtc	ctcctggggc	agaaagaagc	3960
aggcacatcc	ccttctctgt	gacacaccct	gtccacgccc	ctgggttctta	gttccagccc	4020
cactcatagg	acactcatag	ctcaggaggg	ctccgccttc	aatcccaccc	gctaaagtac	4080

```

ttggagcggg ctctccctcc ctcacagcc caccaaacca aacctagcct ccaagagtgg 4140
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4200
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4260
ttaatcttcc gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4320
ggatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4380
aaagaacatg tgagcaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct 4440
ggcgttttcc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4500
gaggtggcga aaccgcagag gactataaag ataccaggcg tttccccctg gaagctccct 4560
cgtgcgctct cctgttccga ccctgcgcgt taccggatag ctgtccgcct ttctcccttc 4620
gggaagcgtg gcgctttctc atagctcacg ctgtagggtat ctgagttcgg tgtaggtcgt 4680
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata 4740
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4800
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4860
gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 4920
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4980
cgggtggttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 5040
tcctttgatc tttctacgg ggtctgacgc tcagtggaaac gaaaactcac gttaagggat 5100
tttggtcagt agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5160
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5220
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5280
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5340
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttggttga 5400
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5460
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5520
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5580
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5640
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5700
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5760
aatttccctc cgtcaaaaat aagggtatca agtgagaaat caccatgagt gacgactgaa 5820
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5880
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5940
tgagcgagac gaaatacgcg atcgcgtgtta aaaggacaat tacaaacagg aatcgaatgc 6000
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6060
tctaatacct ggaatgctgt tttcccgggg atcgcagtggt tgagtaacca tgcacatca 6120
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6180
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6240
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6300
tcgcgagccc atttataccc atataaatca gcatccatgt tggaaattta tcgcggcctc 6360
gagcaagacg tttcccggtg aatatggctc ataacacccc ttgtattact gtttatgtaa 6420
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6480
ttttgagaca caacgtggct tttccccccc cccattattt gaagcattta tcagggttat 6540
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6600
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6660
acctataaaa ataggcggtat caccaggccc tttcgtc 6697

```

&lt;210&gt; 22

&lt;211&gt; 6712

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 22

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt battaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcca accgaccccc ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaggt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatltatta tttacaaatt cacatataca acaacgcgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctcggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagctcag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860
tgcatcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccc 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccagcct gcgtgcccac cgaccccaac cccagggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgacagagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280
ccctgtgctg gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340
gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcaac atcagacca 2400
gcacccgcgg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 2460
tcgacaacga caccaccagc tacagcctga ccagctgcaa caccagcgtg atcaccagg 2520
cctgccccaa ggtgagcttc gagcccatcc ccaaccacta ctgcgcccc cccggcttcg 2580
ccatcctgaa gtgcaaggac aagaagtcca acggcaaggg cccctgcacc aacgtgagca 2640
ccgtgcagtg caccacggc atccgccccg tgggtgagcac ccagctgctg ctgaacggta 2700
gcctggccga ggaggaggtg gtgatccgca gcgctaactt cgccgacaac gccaaaggta 2760
tcacgtgca gctgaacgag agcgtggaga tcaactgcac ccgcgctagc gccactgca 2820
acctgagccg cgccaagtgg aacgacaccc tgaacaagat cgtgatcaag ctgcgcgagc 2880
agttcggcaa caagaccatc gtgttcaagc acagcagcgg cggcgacccc gagatcgtga 2940
cccacagctt caactgcggc ggcgagttct tctactgcaa cagcaccag ctgttcaaca 3000
gcacctggtt caacagcacc tggagcaccg agggcagcaa caacaccgag ggcagcgaca 3060
ccatcacctt gccctgccgc atcaagcaga tcatcaacat gtggcagaag gtgggcaagg 3120
ccatgtacgc ccccccatc agcggccaga tccgtgcag cagcaacatc accggcctgc 3180
tgctgacctg cgacggcggc aacagcaaca acgagagcga gatcttccgt cggggcggcg 3240

```



gcgacatgcg	cgacaactgg	cgcagcgcgc	tgtacaagta	caaggtgggtg	aagatcgagc	3300
ccctgggcgt	ggccccacc	aaggccaagc	ttaccgtcca	ggcccgccag	ctgctgagcg	3360
gcacgtgca	gcagcagaac	aacctgctgc	gcgccatcga	ggcccagcag	cacctgctgc	3420
agctgaccgt	gtggggcatc	aagcagctgc	aggcccgcac	cctggccgtg	gagcgctacc	3480
tgaaggacca	gcagcagctc	gagcagatct	ggaaccacac	cacctggatg	gagtgggacc	3540
gcgagatcaa	caactacacc	agcctgatcc	acagcctgat	cgaggagagc	cagaaccagc	3600
aggagaagaa	cgagcaggag	ctgctggagc	tggacaagtg	ggccagcctg	tggaaactggt	3660
tcaacatcac	caactggctg	tgggtgaggat	ccagatctgc	tgtgccttct	agttgccagc	3720
catctgttgt	ttgcccctcc	cccgtgcctt	ccttgaccct	ggaagggtgc	actcccactg	3780
tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	3840
tgggggggtg	ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	3900
ctggggatgc	ggtgggctct	atgggtaccc	aggtgctgaa	gaattgacct	ggttcctcct	3960
gggccagaaa	gaagcaggca	catccccttc	tctgtgacac	accctgtcca	cgcccctggt	4020
tcttagttcc	agccccactc	ataggacact	catagctcag	gagggctccg	ccttcaatcc	4080
caccgcctaa	agtacttgga	gcggtctctc	cctccctcat	cagcccacca	aaccaaacct	4140
agcctccaag	agtgggaaga	aattaaagca	agataggcta	ttaagtgcag	agggagagaa	4200
aatgcctcca	acatgtgagg	aagtaatgag	agaaatcata	gaattttaag	gccatgattt	4260
aaggccatca	tggccttaat	cttccgcttc	ctcgctcact	gactcgctgc	gctcggtcgt	4320
tccggtgcgg	cgagcgggat	cagctcactc	aaaggcggta	atacggttat	ccacagaatc	4380
aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	4440
aaaggccgcg	ttgctggcgt	ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	4500
tccgacgtca	agtcagaggt	ggcgaaaccc	gacaggacta	taaagatacc	aggcgtttcc	4560
ccctggaagc	tccctcgtgc	gctctcctgt	tccgaccctg	ccgcttaccg	gatacctgtc	4620
cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatage	tcacgctgta	ggtatctcag	4680
ttcggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	gaaccccccg	ttcagcccga	4740
ccgctgcgcc	ttatccggta	actatcgtct	tgagtccaac	ccggtaaagc	acgacttate	4800
gccactggca	gcagccactg	gtaacaggat	tagcagagcg	aggtatgtag	gcggtgctac	4860
agagttcttg	aagtgggtggc	ctaactacgg	ctacactaga	agaacagtat	ttggtatctg	4920
cgctctgctg	aagccagtta	ccttcggaaa	aagagttggt	agctcttgat	ccggcaaaca	4980
aaccaccgct	ggtagcgggtg	gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	5040
aggatctcaa	gaagatcctt	tgatcttttc	tacggggctc	gacgctcagt	ggaacgaaaa	5100
ctcacgttaa	gggatttttg	tcatgagatt	atcaaaaagg	atcttcacct	agatcctttt	5160
aaabtaaaaa	tgaagtthta	aatcaatcta	aagtatatat	gagtaaactt	ggtctgacag	5220
ttaccaatgc	ttaatcagtg	aggcacctat	ctcagcgatc	tgtctatttc	gttcatccat	5280
agttgcctga	ctcggggggg	gggggcgctg	aggtctgcct	cgtgaagaag	gtgttgctga	5340
ctcataccag	gcctgaatcg	ccccatcctc	cagccagaaa	gtgaggggagc	cacgggttgat	5400
gagagctttg	ttgtagggtg	accagttggt	gatttttgaa	ttttgctttg	ccacggaacg	5460
gtctgcgttg	tccgggaagat	gcgtgatctg	atccttcaac	tcagcaaaaag	ttcgatttat	5520
tcaacaaagc	cgccgtcccc	tcaagtcagc	gtaatgctct	gccagtgtta	caaccaatta	5580
accaattctg	attagaaaaa	ctcatcgagc	atcaaatgaa	actgcaattt	attcatatca	5640
ggattatcaa	taccatattt	ttgaaaaagc	cgtttctgta	atgaaggaga	aaactcacccg	5700
aggcagttcc	ataggatggc	aagatcctgg	tatcggctctg	cgattccgac	tccgtccaaca	5760
tcaatacaac	ctattaattt	cccctcgtca	aaaataagggt	tatcaagtga	gaaatcacca	5820
tgagtgcaga	ctgaatccgg	tgagaatggc	aaaagcttat	gcatttcttt	ccagacttgt	5880
tcaacaggcc	agccattacg	ctcgtcatca	aaatcactcg	catcaaccaa	accgttatte	5940
attcgtgatt	gcgcctgagc	gagacgaaat	acgcgatcgc	tgttaaaagg	acaattacaa	6000
acaggaatcg	aatgcaaccg	gcgcaggaa	actgccagcg	catcaacaat	attttcacct	6060
gaatcaggat	attcttctaa	tacctggaat	gctgttttcc	cggggatcgc	agtgggtgagt	6120
aaccatgcat	catcaggagt	acggataaaa	tgcttgatgg	tcggaagagg	cataaattcc	6180
gtcagccagt	ttagtctgac	catctcatct	gtaacatcat	tggcaacgct	acctttgcca	6240
tgtttcagaa	acaactctgg	cgcacggggc	ttcccataca	atcgatagat	tgtcgcacct	6300
gattgcccga	cattatcgcg	agccatttta	taccatata	aatcagcatc	catgttggaa	6360
tttaatcgcg	gcctcgagca	agacgtttcc	cgttgaatat	ggctcataac	accccttgta	6420
ttactgttta	tgtaagcaga	cagttttatt	gttcatgatg	atatattttt	atcttgtgca	6480
atgtaacatc	agagattttg	agacacaacg	tggctttccc	cccccccca	ttattgaage	6540

atztatcagg gttattgtct catgagcgga tacatatattg aatgtattta gaaaaataaa 6600  
 caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 6660  
 attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tc 6712

<210> 23

<211> 6715

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 23

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
 ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
 cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
 agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140  
 taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
 tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260  
 tattggctat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320  
 ggatggggtc ccattttatta tttaaaaatt cacatatata acaacgccgt cccccgtgcc 1380  
 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500  
 agcggctcat ggtcgtctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga atgcgcgtga 1920  
 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980  
 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040  
 tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100  
 aggtgcacaa cgtgtgggac acccagcct gcgtgcccac cgaccccaac ccccaggagg 2160  
 tgggtgctgt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220  
 tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280  
 ccctgtgctg gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340  
 gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcaac atcagacca 2400

gcatccgcgg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 2460  
tcgacaacga caccaccagc tacagcctga ccagctgcaa caccagcgtg atcaccacagg 2520  
cctgccccaa ggtgagcttc gagcccatcc ccaaccacta ctgcgcccccc gccggcttcg 2580  
ccatcctgaa gtgcaaggac aagaagttca acggcaaggg cccctgcacc aacgtgagca 2640  
ccgtgcagtg caccacaggc atccgccccg tggtagcac ccagctgctg ctgaacggta 2700  
gcctggccga ggaggagggtg gtgatccgca gcgctaactt cgccgacaac gccaagggtga 2760  
tcatcgtgca gctgaacgag agcgtggaga tcaactgcac ccgccccaac aacaacacccc 2820  
gcaagagcat ccacatcggc cccggccgcg ccttctacac caccggcgag atcatcggcg 2880  
acatccgcca ggcccactgc aacctgagcc gcgccaagtg gaacgacacc ctgaacaaga 2940  
tcgtgatcaa gctgcgcgag cagttcggca acaagaccat cgtgttcaag cacagcagcg 3000  
gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc ttctactgca 3060  
acagcgctag cctgcccctgc cgcataaagc agatcatcaa catgtggcag aagggtgggca 3120  
aggccatgta cgcgcccccc atcagcggcc agatccgctg cagcagcaac atcaccggcc 3180  
tgctgctgac ccgcgacggc ggcaacagca acaacgagag cgagatcttc cgtccggggcg 3240  
gcggcgacat gcgcgacaac tggcgagcg agctgtacaa gtacaagggtg gtgaagatcg 3300  
agccccctggg cgtggcccccc accaaggcca agcttaccgt ccaggccccgc cagctgctga 3360  
gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag cagcacctgc 3420  
tgcagctgac cgtgtggggc atcaagcagc tgcaggccccg caccctggcc gtggagcgct 3480  
acctgaagga ccagcagcag ctcgagcaga tctggaacca caccacctgg atggagtggg 3540  
accgcgagat caacaactac accagcctga tccacagcct gatcgaggag agccagaacc 3600  
agcaggagaa gaacgagcag gagctgctgg agctggacaa gtggggccagc ctgtggaact 3660  
ggttcaacat caccaactgg ctgtggtgag gatccagatc tgctgtgcct tctagttgcc 3720  
agccatctgt tgtttgcccc tccccctgct ctcccttgac cctggaagggt gccactccca 3780  
ctgtcctttc ctaataaaaat gaggaaattg catcgcatcg tctgagtagg tgtcattcta 3840  
ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac aatagcaggc 3900  
atgctgggga tgcgggtgggc tctatgggta cccaggtgct gaagaattga cccggttcct 3960  
cctggggccag aaagaagcag gcacatcccc ttctctgtga cacacctgt ccacgccccct 4020  
ggttcttagt tccagcccca ctcataggac actcatagct caggagggct ccgccttcaa 4080  
tcccaccgc taaagtactt ggagcggctc ctccctccct catcagccca ccaaaccaaa 4140  
cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg cagagggaga 4200  
gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt aaggccatga 4260  
tttaaggcca tcatggcctt aatcttccgc ttccctcgctc actgactcgc tgcgctcggt 4320  
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgggt tatccacaga 4380  
atcaggggat aacgcaggaa agaacatgtg agcaaaaaggc cagcaaaaagg ccaggaaccg 4440  
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca 4500  
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 4560  
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgcccgtta ccggatacct 4620  
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 4680  
cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 4740  
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccgggtaa gacacgactt 4800  
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgagggtat taggcgggtg 4860  
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat 4920  
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 4980  
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 5040  
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 5100  
aaactcacgt taagggattt tggctcatgag attatcaaaa aggatcttca cctagatcct 5160  
tttaaatata aaatgaagtt tbaaatcaat ctaaagtata tatgagtaaa cttgggtctga 5220  
cagttaccaaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttctgtctatc 5280  
catagttgcc tgactcgggg gggggggggc ctgaggctcg cctcgtgaag aagggtgtgc 5340  
tgactcatal caggcctgaa tcgccccatc atccagccag aaagtgaggg agccacgggt 5400  
gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct ttgccacgga 5460  
acggtctgcg ttgtcgggaa gatgcgtgat ctgaccttc aactcagcaa aagttcgatt 5520  
tattcaacaa agccgcgctc ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa 5580  
ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa ttattcata 5640  
tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca 5700



```

ccgaggcagt tccataggat ggcaagatcc tggatatcgg ctgcgattcc gactcgtcca 5760
acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag tgagaaatca 5820
ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc tttccagact 5880
tggtcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaaccgta 5940
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 6000
caaacaggaa tcgaatgcaa ccggcgcagg aacactgcc aacactgcc aatattttca 6060
cctgaatcag gatattcttc taataacctg aatgctgttt tcccggggat cgcagtgggtg 6120
agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag aggcataaat 6180
tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac gctacctttg 6240
ccatgtttca gaaacaactc tggcgcacgc ggcttcccat acaatcgata gattgtcgca 6300
cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc atccatgttg 6360
gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat aacacccctt 6420
gtattactgt ttatgtaagc agacagtttt attgttcatt atgatatatt tttatcttgt 6480
gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc ccattattga 6540
agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 6600
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6660
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtc 6715

```

&lt;210&gt; 24

&lt;211&gt; 6619

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 24

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggtcctatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggcgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctg cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560

```

```

agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggcttttcc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccc 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccacgcct gcgtgcccac cgaccccaac ccccaggagg 2160
tgggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgctg aagctgaccc 2280
ccctgtgctg gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaggcctg cccaagggtg agcttcgagc ccatcccaaa ccactactgc gccccgcgcg 2400
gcttcgcat cctgaagtgc aaggacaaga agttcaacgg caaggggccc tgcaccaacg 2460
tgagcacctg gcagtgcacc cacggcatcc gcccgtggt gagcaccag ctgctgctga 2520
acggtagcct ggccgaggag gaggtggtga tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgcg cccaacaaca 2640
acacccgcaa gagcatccac atcggccccg gccgcgcctt ctacaccacc ggcgagatca 2700
tcggcgacat ccgccaggcc cactgcaacc tgagccgcgc caagtggaa gacaccctga 2760
acaagatcgt gatcaagctg cgcgagcagt tcggcaacaa gaccatcgtg ttcaagcaca 2820
gcagcgggcg cgaccccgag atcgtgaccc acagcttcaa ctgcggcggc gagttcttct 2880
actgcaacag caccagctg ttcaacagca cctggttcaa cagcacctgg agcaccgagg 2940
gcagcaacaa caccgagggc agcgacacca tcaccctgcc ctgccgcctc aagcagatca 3000
tcaacatgtg gcagaagggt ggcaaggcca tgtacgcccc ccccatcagc ggccagatcc 3060
gctgcagcag caacatcacc ggctgctgc tgaccgcga cggcggaac agcaacaacg 3120
agagcgagat cttccgtccg ggcgcgcgcg acatgcgcga caactggcgc agcgagctgt 3180
acaagtacaa ggtggtgaag atcgagcccc tggcgctggc cccaccaag gccaaagctta 3240
ccgtccaggc ccgccagctg ctgagcgcca tcgtgcagca gcagaacaa ctgctgcgcg 3300
ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg 3360
cccgcaccct ggccgtggag cgctacctga aggaccagca gcagctcgag cagatctgga 3420
accacaccac ctggatggag tgggaccgcg agatcaacaa ctacaccagc ctgatccaca 3480
gcctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg ctggagctgg 3540
acaagtgggc cagcctgtgg aactggttca acatcaccaa ctggctgtgg tgaggatcca 3600
gatctgctgt gccttctagt tgccagccat ctgttgtttg cccctcccc gtgccttctt 3660
tgaccctgga aggtgccact cccactgtcc ttctctaata aaatgaggaa attgcatcgc 3720
attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac agcaaggggg 3780
aggattggga agacaatagc aggcattgct gggatgcggg gggctctatg ggtaccagg 3840
tgctgaagaa ttgaccgggt tcctcctggg ccagaaagaa gcaggcacat ccccttctct 3900
gtgacacacc ctgtccacgc cctggttct tagttccagc cccactcata ggacactcat 3960
agctcaggag ggctccgcct tcaatccac ccgctaaagt acttgagcg gtctctccct 4020
ccctcatcag cccaccaaac caaacctagc ctccaagagt gggagaaat taaagcaaga 4080
taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 4140
aatcatagaa btttaaggcc atgatttaag gccatcatgg ccttaatctt ccgcttctct 4200
gctcactgac tcgctgcgct cggctcgttc gctgcggcga gcggtatcag ctactcaaa 4260
ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4320
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 4380
ccgccccctt gacgagcatc aaaaaaatcg acgctcaagt cagagggtggc gaaacccgac 4440
aggactataa agataccagg cgtttccccc tggaaagctc ctctgctgct ctctgttcc 4500
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttct 4560
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg 4620
tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggttaact atcgtcttga 4680
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4740
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta 4800
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 4860

```

```

agttggtagc tcttgatccg gcaaacaaac caccgctggt agcgggtggtt tttttggttg 4920
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 4980
ggggtctgac gctcagtga acgaaaactc acgttaaggg attttgggtca tgagattatc 5040
aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agttttaaat caatctaaag 5100
tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc 5160
agcgatctgt ctatttcgtt catccatagt tgcctgactc gggggggggg ggcgctgagg 5220
tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag 5280
ccagaaagtg agggagccac gggtgatgag agctttgttg taggtggacc agttgggtgat 5340
tttgaacttt tgctttgccg cggaacggtc tgcgttgctg ggaagatgcg tgatctgac 5400
cttcaactca gcaaaagtgc gatttattca acaaagccgc cgtcccgtca agtcagcgta 5460
atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc atcgagcatc 5520
aatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5580
ttctgtaatg aaggagaaaa ctcccgagg cagttccata ggatggcaag atcctgggat 5640
cggctctgca ttccgactcg tccaacatca atacaacctt ttaatttccc ctgctcaaaa 5700
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccgggtg gaatggcaaa 5760
agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc gtcataaaaa 5820
tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg 5880
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 5940
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6000
gttttcccgg ggatcgcatg ggtgagtaac catgcatcat caggagtacg gataaaatgc 6060
ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcatctgta 6120
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6180
ccatacaatc gatagattgt cgcacctgat tgcctgacat tatcgcgagc ccatttatac 6240
ccatataaat cagcatccat gttggaattt aatcgcgccg tcgagcaaga cgtttcccgt 6300
tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattgtt 6360
catgatgata tatttttata ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6420
ctttccccc ccccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatag 6480
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaaa 6540
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6600
atcacgaggc cctttcgtc                                     6619

```

&lt;210&gt; 25

&lt;211&gt; 6655

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 25

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgccttggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780

```



```

cgtcaatggg agtttgTTTT ggcacaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcggtgtacgg tgggaggtct atataagcag 900
agctcgTTta gtgaaccgtc agatcgccctg gagacgccat ccacgtgttt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgtacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggtttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgtc acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt abtaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtggggc acccacgcct gcgtgcccac cgaccccaac cccagaggag 2160
tggtgctggt gaacgtgacc gagaacttct acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaagggtg cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 2460
aggcctgccc caagggtgagc ttcgagccca tcccaacca ctactgcgc cccgcgggct 2520
tcgccatcct gaagtgcag gacaagaagt tcaacggcaa gggccctgc accaacgtga 2580
gcaccgtgca gtgcaccac ggcatccgcc ccgtgggtgag caccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcgct agcgcctact 2760
gcaacctgag ccgcgccaag tggaaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2820
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2880
tgaccacag cttcaactgc ggcggcgagt tcttctactg caacagcacc cagctgttca 2940
acagcacctg gttcaacagc acctggagca ccgagggcag caacaacacc gagggcagcg 3000
acaccatcac cctgccctgc cgcataaagc agatcatcaa catgtggcag aaggtgggca 3060
aggccatgta cgcgcgcgc atcagcggcc agatccgctg cagcagcaac atcaccggcc 3120
tgctgctgac ccgcgacggc ggcaacagca acaacgagag cgagatcttc cgtccgggcg 3180
gcggcgacat gcgcgacaa tggcgacgcg agctgtacaa gtacaagggtg gtgaagatcg 3240
agccctggg cgtggccccc accaaggcca agcttaccgt ccaggcccg cagctgctga 3300
gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag cagcacctgc 3360
tgacgtgac cgtgtggggc atcaagcagc tgcaggcccg caccctggcc gtggagcgct 3420
acctgaagga ccagcagcag ctcgagcaga tctggaacca caccacctgg atggagtggg 3480
accgcgagat caacaactac accagcctga tccacagcct gatcgaggag agccagaacc 3540
agcaggagaa gaacgagcag gagctgctgg agctggacaa gtgggcccagc ctgtggaact 3600
ggttcaacat caccaactgg ctgtggtgag gatccagatc tgctgtgcct tctagttgcc 3660
agccatctgt tgtttgcccc tccccgtgc ctctcttgac cctggaagggt gccactccca 3720
ctgtcctttc ctaataaaat gaggaattg catcgcattg tctgagtagg tgtcatctca 3780
ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac aatagcaggc 3840
atgctgggga tgcggtgggc tctatgggta cccaggtgct gaagaattga cccggttcc 3900
cctgggccag aaagaagcag gcacatcccc ttctctgtga cacacctgt ccacgcccct 3960
ggttcttagt tccagcccca ctcataggac actcatagct caggagggct ccgccttcaa 4020
tcccaccgc taaagtactt ggagcgggtc ctccctccct catcagccca ccaaaccaaa 4080

```



cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg cagagggaga 4140  
gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt aaggccatga 4200  
tttaaggcca tcatggcctt aatcttccgc ttcctcgtc actgactcgc tgcgctcgg 4260  
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 4320  
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 4380  
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca 4440  
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 4500  
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 4560  
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 4620  
cagttcgggtg taggtcgctt gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 4680  
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtta gacacgactt 4740  
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 4800  
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat 4860  
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 4920  
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcaaaa 4980  
aaaaggatct caagaagatc ctttgatctt ttctacggg tctgacgctc agtggaacga 5040  
aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct 5100  
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga 5160  
cagttaccaaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 5220  
catagttgcc tgactcgggg gggggggcg ctgaggtctg cctcgtgaag aagggtgtgc 5280  
tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg agccacggtt 5340  
gatgagagct ttgtttagg tggaccagtt ggtgattttg aacttttgct ttgccacgga 5400  
acggtctgcg ttgtcgggaa gatgcgtgat ctgaccttc aactcagcaa aagttcgatt 5460  
tattcaacaa agccgccgtc ccgtcaagtc agcgtaatgc tctgccagtg ttacaacca 5520  
ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa ttattcata 5580  
tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca 5640  
ccgaggcagt tccataggat ggcaagatcc tggatcggg ctgcatcacc gactcgtcca 5700  
acatcaatac aacctattaa tttccctcgc tcaaaaataa gggtatcaag tgagaaatca 5760  
ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc tttccagact 5820  
tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaacggtta 5880  
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 5940  
caaacaggaa tcgaatgcaa ccggcgagc aacactgcca gcgcatcaac aatattttca 6000  
cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat cgcagtgggtg 6060  
agtaaccatg catcatcagg agtacggata aaatgcttga tggcgggaag aggcataaat 6120  
tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac gctacctttg 6180  
ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca 6240  
cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc atccatgttg 6300  
gaatttaatc gcggcctcga gcaagacggt tcccgttgaa tatggctcat aacacccctt 6360  
gtattactgt ttatgtaagc agacagtttt attgttcatg atgatataat tttatcttgt 6420  
gcaatgtaac atcagagatt ttgagacaca acgtggcctt ccccccccc ccattattga 6480  
agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 6540  
aaacaaatag ggggtccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6600  
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtc 6655

&lt;210&gt; 26

&lt;211&gt; 6658

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 26

tcgcgcggttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	gagacgggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
cccgcctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgcc	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttgucagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgta	780
cgtcaatggg	agtttggttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgcttg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggt	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtcc	1380
cgcagttttt	attaaacata	gcgtgggato	tccacgcgaa	tctcggttac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggt	ccatgcctcc	1500
agcggctcat	ggtcgtcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgceg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcag	gtagtctgag	cagtactcgt	tgctgcgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgctggg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	acccacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttct	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgctg	gagcctgaag	tgcaccgacg	ctagcaagaa	ctgcagcttc	aacatcagca	2340
ccagcatccg	cggcaagggt	cagaaggagt	acgccttctt	ctacaagctg	gacatcatcc	2400
ccatcgacaa	cgacaccacc	agctacagcc	tgaccagctg	caacaccagc	gtgatcacc	2460
aggcctgccc	caagggtgag	ttcgagccca	tccccaaaca	ctactgcgcc	cccgccgggt	2520
tcgccatcct	gaagtgcag	gacaagaagt	tcaacggcaa	gggcccctgc	accaacgtga	2580
gcaccgtgca	gtgcaccac	ggcatccgcc	ccgtgggtgag	caccagctg	ctgctgaacg	2640
gtagcctggc	cgaggaggag	gtggtgatcc	gcagcgctaa	cttcgcccag	aacgccaagg	2700
tgatcatcgt	gcagctgaac	gagagcgtgg	agatcaactg	caccgcctcc	aacaacaaca	2760
cccgaagag	catccacatc	ggccccggcc	gcgccttcta	caccaccggc	gagatcatcg	2820
gcgacatccg	ccaggccccc	tgaacctga	gccgcgcaa	gtggaacgac	accctgaaca	2880
agatcgtgat	caagctgcgc	gagcagttct	gcaacaagac	catcgtgttc	aagcacagca	2940
gcggcggcga	ccccgagatc	gtgaccacac	gcttcaactg	cggcggcgag	ttcttctact	3000
gcaacagcgc	tagcctgccc	tggcgcatca	agcagatcat	caacatgtgg	cagaagggtg	3060
gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	ctgcagcagc	aacatcaccg	3120
gcctgctgct	gacccgcgac	ggcggcaaca	gcaacaacga	gagcgagatc	ttccgtccgg	3180
gcggcggcga	catgcgcgac	aactggcgca	gcgagctgta	caagtacaag	gtggtgaaga	3240
tcgagccct	gggcgtggcc	cccaccaagg	ccaagcttac	cgtccaggcc	cgccagctgc	3300

tgagcggcat	cgtgcagcag	cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	3360
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcaggc	ccgcaccctg	gccgtggagc	3420
gctacctgaa	ggaccagcag	cagctcgagc	agatctggaa	ccacaccacc	tggatggagt	3480
gggaccgcga	gatcaacaac	tacaccagcc	tgatccacag	cctgatcgag	gagagccaga	3540
accagcagga	gaagaacgag	caggagctgc	tggagctgga	caagtgggcc	agcctgtgga	3600
actggttcaa	catcaccaac	tggctgtggt	gaggatccag	atctgctgtg	ccttctagtt	3660
gccageccatc	tgttgtttgc	ccctcccccg	tgcccttcctt	gaccctggaa	ggtgccactc	3720
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3780
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3840
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccaggt	gctgaagaat	tgaccgggtt	3900
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	3960
cctggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	4020
caatcccacc	cgctaaagta	cttggagcgg	tctctccctc	cctcatcagc	ccaccaaacc	4080
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	4140
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	4200
tgatttaagg	ccatcatggc	cttaatcttc	cgcttctctg	ctcactgact	cgctgcgctc	4260
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaa	gcggtaatat	ggttatccac	4320
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4380
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgccccctg	acgagcatca	4440
caaaaatcga	cgctcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	4500
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	4560
cctgtccgcc	tttctccctt	cggaagcgt	ggcgctttct	catagctcac	gctgtaggta	4620
tctcagttcg	gtgtaggctg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	4680
gcccagaccgc	tgcgccttat	ccggtaaacta	tcgtcttgag	tccaaccggg	taagacacga	4740
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	4800
tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	4860
tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggttagct	cttgatccgg	4920
caaaacaacc	accgctggta	gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	4980
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	5040
cgaaaactca	cgttaaggga	ttttggtcat	gagattatca	aaaaggatct	tcacctagat	5100
cctttttaa	taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	5160
tgacagttac	caatgcttaa	tcagttaggc	acctatctca	gcgatctgtc	tatttcgttc	5220
atccatagtt	gcctgactcg	gggggggggg	gcgctgaggt	ctgcctcggt	aagaagggtg	5280
tgctgactca	taccaggcct	gaatcgcccc	atcatccagc	cagaaagtga	gggagccacg	5340
gttgatgaga	gctttgttgt	aggtggacca	gttggtgatt	ttgaactttt	gctttgccac	5400
ggaacggtct	gcgttgctcg	gaagatgcgt	gatctgatcc	ttcaactcag	caaaagtctg	5460
atthattcaa	caaagccgcc	gtcccgctca	gtcagcgtaa	tgctctgcca	gtgttacaac	5520
caattaacca	attctgatta	gaaaaactca	tcgagcatca	aatgaaactg	caatttattc	5580
atatcaggat	tatcaatacc	atatttttga	aaaagccggt	tctgtaatga	aggagaaaac	5640
tcaccgaggg	agttccatag	gatggcaaga	tcctgggtatc	ggctctgcgat	tccgactcgt	5700
ccaacatcaa	tacaacctat	taatttcccc	tcgtcaaaaa	taagggttatc	aagtgagaaa	5760
tcaccatgag	tgacgactga	atccggtgag	aatggcaaaa	gcttatgcat	ttctttccag	5820
acttggtcaa	caggccagcc	attacgctcg	tcatcaaaat	cactcgcatc	aaccaaacccg	5880
ttattcattc	gtgattgcgc	ctgagcgaga	cgaaatacgc	gatcgctgtt	aaaaggacaa	5940
ttacaaacag	gaatcgaatg	caaccggcgc	aggaacactg	ccagcgcatc	aacaatatatt	6000
tcacctgaat	caggatattc	ttctaatacc	tggaaatgctg	ttttcccggg	gatcgcagtg	6060
gtgagtaacc	atgcatactc	aggagtacgg	ataaaatgct	tgatggtcgg	aagaggcata	6120
aattccgtca	gccagttag	tctgaccatc	tcactctgtaa	catcattggc	aacgctacct	6180
ttgccatggt	tcagaaacaa	ctctggcgca	tcgggcttcc	catacaatcg	atagattgtc	6240
gcacctgatt	gcccagacatt	atcgcgagcc	catttatacc	catataaatc	agcatccatg	6300
ttggaattta	atcgcgccct	cgagcaagac	gtttcccggt	gaatatggct	cataaacacc	6360
cttgatttac	tgtttatgta	agcagacagt	tttattgttc	atgatgatat	atthttatct	6420
tgtgcaatgt	aacatcagag	atthttgagac	acaacgtggc	tttccccccc	ccccatttat	6480
tgaagcattt	atcagggtta	ttgtctcatg	agcgataaca	tatttgaatg	tatttagaaa	6540
aataaaca	taggggttcc	gcgcacattt	ccccgaaaag	tgccacctga	cgtctaagaa	6600



accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6658

<210> 27

<211> 6616

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 27

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttggttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcac gctatactgt ttttggtctg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggtctt tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
agggtgcaca cgtgtgggac acccagcct gcgtgcccac cgaccccaac cccagggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280
ccctgtgcgt gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340
gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcgct agcaccagct 2400
gcaacaccag cgtgatcacc caggcctgcc ccaaggtgag cttcgagccc atccccacc 2460
actactgcgc ccccgccggc ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca 2520

```

```

agggccctg caccaacgtg agcaccgtgc agtgcaccca cggcatccgc cccgtggtga 2580
gcaccagct gctgctgaac ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta 2640
acttcgccga caacgccaaag gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact 2700
gcaccgcgc tagcgcacac tgcaacctga gccgcgcaa gtggaacgac accctgaaca 2760
agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc aagcacagca 2820
gcggcggcga ccccgagatc gtgaccaca gcttcaactg cggcggcgag ttcttctact 2880
gcaacagcac ccagctgttc aacagcacct ggttcaacag cacctggagc accgagggca 2940
gcaacaacac cgagggcagc gacaccatca ccctgccctg ccgcatcaag cagatcatca 3000
acatgtggca gaaggtgggc aaggccatgt acgccccccc catcagcggc cagatccgct 3060
gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaacagc aacaacgaga 3120
gcgagatctt ccgtccgggc ggcggcgaca tgcgcgaaa ctggcgcagc gagctgtaca 3180
agtacaagggt ggtgaagatc gagccccctg gcgtggcccc caccaaggcc aagcttaccg 3240
tccaggcccc ccagctgctg agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca 3300
tcgaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc 3360
gcaccctggc cgtggagcgc tacctgaagg accagcagca gctcgagcag atctggaacc 3420
acaccacctg gatggagtgg gaccgcgaga tcaacaacta caccagcctg atccacagcc 3480
tgatcgagga gagccagaac cagcaggaga agaacgagca ggagctgctg gagctggaca 3540
agtgggccag cctgtggaac tggttcaaca tcaccaactg gctgtggtga ggatccagat 3600
ctgctgtgcc ttctagtgtc cagccatctg ttgtttgcc ctccccctg ccttccttga 3660
ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcat 3720
gtctgagtag gtgtcattct attctggggg gtgggggtgg gcaggacagc aagggggagg 3780
attgggaaga caatagcagg catgctgggg atgcgggtgg ctctatgggt acccaggtgc 3840
tgaagaattg acccggttcc tctgggcca gaaagaagca ggcacatccc cttctctgtg 3900
acacacctg tccacgcccc tggttcttag ttccagcccc actcatagga cactcatagc 3960
tcaggagggc tccgccttca atcccaccg ctaaagtact tggagcggc tctccctccc 4020
tcatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4080
gctattaagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4140
catagaattt taaggccatg atttaaggcc atcatggcct taatcttccg cttcctcgct 4200
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 4260
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4320
ccagcaaaag gccaggaacc gtaaaaagge cgcgttgctg gcgtttttcc ataggctccg 4380
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4440
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 4500
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4560
tagctcacgc tgtaggatc tcagttcggg ttaggtcggt cgctccaagc tgggctgtgt 4620
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4680
caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4740
agcgaggatg gtaggcgggt ctacagagtt cttgaagtgg tggcctaact acggctacac 4800
tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 4860
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 4920
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 4980
gtctgacgct cagtggaaacg aaaactcacg ttaagggtatt ttggtcatga gattatcaaa 5040
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 5100
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5160
gatctgtcta ttctgttcat ccatagttgc ctgactcggg gggggggggc gctgaggtct 5220
gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga atcgccccat catccagcca 5280
gaaagtgagg gagccacggt tgatgagagc tttgtttag gtggaccagt tgggtgatttt 5340
gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga tctgacctt 5400
caactcagca aaagtctgat ttattcaaca aagccgcgct ccgtcaagt cagcgtaatg 5460
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa 5520
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgttct 5580
tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5640
tctgcgattc cgactcgtec aacatcaata caacctatta atttccctc gtcaaaaata 5700
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5760
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5820

```

```

ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 5880
tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgagc gaacactgcc 5940
agcgcatcaa caatattttc acctgaatca ggatattcct ctaataacctg gaatgctgtt 6000
ttcccgggga tcgcagtggg gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6060
atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6120
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca 6180
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca 6240
tataaatcag catccatgtt ggaatttaac cgcggcctcg agcaagacgt ttcccgttga 6300
atatggctca taacacccct tgtattactg tttatgtaag cagacagttt tattgttcat 6360
gatgatata ttttatcttg tgcaatgtaa catcagagat tttgagacac aacgtggctt 6420
tcccccccc cccattattg aagcatttat cagggttatt gtctcatgag cggatacata 6480
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6540
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6600
acgaggccct ttcgtc                                     6616

```

&lt;210&gt; 28

&lt;211&gt; 6619

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 28

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgcc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```



gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
 tgcagtcacc gtgcgcgaca cgtgtgatca gatatcgcg cgcctctaga atgcgcgtga 1920  
 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980  
 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040  
 tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100  
 aggtgcacaa cgtgtgggccc acccacgctt gcgtgcccac cgaccccaac cccagaggag 2160  
 tgggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220  
 tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280  
 ccctgtgcgt gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340  
 gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcgct agcaccagct 2400  
 gcaacaccag cgtgatcacc caggcctgcc ccaaggtagg ctctgagccc atccccaacc 2460  
 actactgcgc ccccgccggc ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca 2520  
 agggcccttg caccaacgtg agcacctgac agtgaccca cggcatccgc cccgtggtga 2580  
 gcacccagct gctgctgaac ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta 2640  
 acttcgccga caacgccaag gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact 2700  
 gcacccgccc caacaacaac acccgcaaga gcatccacat cggccccggc cgcgccttct 2760  
 acaccaccgg cgagatcatc ggcgacatcc gccaggccca ctgcaacctg agccgcgcca 2820  
 agtggaaacga caccctgaac aagatcgtga tcaagctgcg cgagcagttc ggcaacaaga 2880  
 ccacgtgtgt caagcacagc agcggcgggc accccgagat cgtgacccac agcttcaact 2940  
 gcggcgggca gttcttctac tgcaacagcg ctacgctgcc ctgccgcatc aagcagatca 3000  
 tcaacatgtg gcagaagggt ggcaaggcca tgtacgcccc ccccatcagc ggccagatcc 3060  
 gctgcagcag caacatcacc ggctgctgc tgacccgcga cggcggaac agcaacaacg 3120  
 agagcgagat ctcccgctcg ggcggcgggc acatgcgcga caactggcgc agcgagctgt 3180  
 acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc cccaccaag gccagctta 3240  
 ccgtccagga ccgcccagct ctgagcggca tgcgtgcagca gcagaacaac ctgctgcgcg 3300  
 ccacgagga ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg 3360  
 cccgcacct ggccgtggag cgctacctga aggaccagca gcagctcgag cagatctgga 3420  
 accacaccac ctggatggag tgggaccgcg agatcaacaa ctacaccagc ctgatccaca 3480  
 gcctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg ctggagctgg 3540  
 acaagtgggc cagcctgtgg aactgggttca acatcaccaa ctggctgtgg tgaggatcca 3600  
 gatctgctgt gccttctagt tgccagccat ctgttgtttg cccctcccc gtgccttct 3660  
 tgaccctgga aggtgccact cccactgtcc ttctctaata aaatgaggaa attgcatcgc 3720  
 attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac agcaaggggg 3780  
 aggattggga agacaatagc aggcattgct gggatgcggg gggctctatg ggtacccagg 3840  
 tgctgaagaa ttgaccgggt tcctcctggg ccagaaagaa gcaggcacat ccccttctct 3900  
 gtgacacacc ctgtccacgc cctgggttct tagttccagc cccactcata ggacactcat 3960  
 agctcaggag ggctccgcct tcaatccccc ccgctaaagt acttgagcgc gtctctccct 4020  
 ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaagaaat taaagcaaga 4080  
 taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 4140  
 aatcatagaa ttttaaggcc atgatttaag gccatcatgg ccttaatctt ccgcttctc 4200  
 gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctactcaaa 4260  
 ggcggtataa cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 4380  
 ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagagggtggc gaaacccgac 4440  
 aggactataa agataccagg cgtttccccc tggaagctcc ctctgctgct ctctgttcc 4500  
 gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttct 4560  
 tcatagctca cgctgtagg atctcagttc ggtgtaggtc gttcgtctca agctgggctg 4620  
 tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 4680  
 gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4740  
 cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta 4800  
 cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 4860  
 agttggtagc tcttgatccg gcaaaaaaac caccgctggg agcgggtgggt tttttgtttg 4920  
 caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 4980  
 ggggtctgac gctcagtggg acgaaaactc acgttaaggg attttggtca tgagattatc 5040



```

aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agtttttaa caatctaaag 5100
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5160
agcgatctgt ctatttcggt catccatagt tgcctgactc gggggggggg ggcgctgagg 5220
tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag 5280
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttgggtgat 5340
tttgaacttt tgctttgcca cggaacggtc tgcgttgctg ggaagatgcg tgatctgate 5400
cttcaactca gcaaaagtgc gatttatcca acaaagccgc cgtcccgtca agtcagcgta 5460
atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc atcgagcate 5520
aatgaaact gcaattttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5580
ttctgtaatg aaggagaaaa ctaccgagg cagttccata ggatggcaag atcctggtat 5640
cggctctgca ttccgactcg tccaacatca atacaacctt ttaattttcc ctcgtcaaaa 5700
ataagggtat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 5760
agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc gtcatacaaa 5820
tactcgcgat caaccaaacc gttattcatt cgtgattgct cctgagcgag acgaaatacg 5880
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 5940
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6000
gttttcccgg ggatcgcatg ggtgagtaac catgcatcat caggagtacg gataaaatgc 6060
ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgacctc ctcatctgta 6120
acatcatttg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6180
ccatacaatc gatagattgt cgcacctgat tgcccgcacat tatcgcgagc ccatttatac 6240
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccgt 6300
tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattgtt 6360
catgatgata tattttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6420
ctttccccc ccccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6480
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa 6540
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6600
atcacgaggc cctttcgtc
6619

```

&lt;210&gt; 29

&lt;211&gt; 6634

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 29

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcatatgc ccagtacatg acctatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaaaaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctc gagacgcat ccacgctgtt ttgacctcca 960

```

tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatc	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttacaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gltgcgggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgcccg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgtcggg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtgggcc	accacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcgt	gatcatggag	aagggcgaga	tcaagaactg	cagcttcaac	atcagcacca	2400
gcatccgcgg	caaggtgcag	aaggagtacg	ccttcttcta	caagctggac	atcatcccca	2460
tcgacaacga	caccaccagc	tacagcctga	ccagctgcaa	caccagcgtg	atcaccagg	2520
cctgccccaa	ggtgagcttc	gagcccatcc	ccaaccacta	ctgcgcccc	gccggcttcg	2580
ccatcctgaa	gtgcaaggac	aagaagtcca	acggcaagg	cccctgcacc	aacgtgagca	2640
ccgtgcagtg	caccacggc	atccgcccc	tgggtgagc	ccagctgctg	ctgaacggta	2700
gcctggccga	ggaggagggt	gtgatccgca	gcgctaactt	cgccgacaac	gccaagggtga	2760
tcacgtgca	gctgaacgag	agcgtggaga	tcaactgcac	ccgcgctagc	gcccactgca	2820
acctgagccg	cgccaagtgg	aacgacaccc	tgaacaagat	cgtgatcaag	ctgcgcgagc	2880
agttcggcaa	caagaccatc	gtgttcaagc	acagcagcgg	cggcgacccc	gagatcgtga	2940
cccacagctt	caactgcggc	ggcgagttct	tctactgcaa	cagcgctagc	ctgccctgcc	3000
gcatcaagca	gatcatcaac	atgtggcaga	aggtgggcaa	ggccatgtac	gccccccca	3060
tcagcggcca	gatccgctgc	agcagcaaca	tcaccggcct	gctgctgacc	cgcgacggcg	3120
gcaacagcaa	caacgagagc	gagatcttcc	gtccgggcgg	cggcgacatg	cgcgacaact	3180
ggcgcagcga	gctgtacaag	tacaagggtg	tgaagatcga	gcccctgggc	gtggcccca	3240
ccaaggccaa	gcttaccgtc	caggcccgc	agctgctgag	cggcatcgtg	cagcagcaga	3300
acaacctgct	gcgcgccatc	gaggcccagc	agcacctgct	gcagctgacc	gtgtggggca	3360
tcaagcagct	gcaggccccg	accctggccg	tggagcgcta	cctgaaggac	cagcagcagc	3420
tcgagcagat	ctggaaccac	accacctgga	tggagtggga	ccgcgagatc	aacaactaca	3480
ccagcctgat	ccacagcctg	atcgaggaga	gccagaacca	gcaggagaag	aacgagcagg	3540
agctgctgga	gctggacaag	tgggccagcc	tgtggaactg	gttcaacatc	accaactggc	3600
tgtggtgagg	atccagatct	gctgtgcctt	ctagtgtcca	gccatctgtt	gtttgcccct	3660
cccccgctgc	ttccttgacc	ctggaagggt	ccactcccac	tgctcctttcc	taataaaatg	3720
aggaaattgc	atcgcattgt	ctgagtaggt	gtcattctat	tctggggggg	ggggtggggc	3780
aggacagcaa	gggggaggat	tgggaagaca	atagcaggca	tgctggggat	gcggtgggct	3840
ctatgggtac	ccaggtgctg	agaattgac	ccggttcctc	ctgggcccaga	aagaagcagg	3900
cacatcccct	tctctgtgac	acacctgtc	cacgcccctg	gttcttagtt	ccagccccac	3960
tcataggaca	ctcatagctc	aggagggctc	cgccttcaat	cccacccgct	aaagtacttg	4020
gagcggctct	tccctccctc	atcagcccac	caaaccacaa	ctagcctcca	agagtgggaa	4080
gaaattaaag	caagataggc	tattaagtgc	agagggagag	aaaatgcctc	caacatgtga	4140
ggaagtaatg	agagaaatca	tagaatttta	aggccatgat	ttaaggccat	catggcctta	4200
atcttccgct	tcctcgcctc	ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggc	4260

```

atcagctcac tcaaaggcgg taatacgggtt atccacagaa tcaggggata acgcaggaaa 4320
gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 4380
gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 4440
gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa gctccctcgt 4500
gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 4560
aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg aggtcgttcg 4620
ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 4680
taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg cagcagccac 4740
tggtaacagg attagcagag cgaggtatgt aggcgggtgt acagagttct tgaagtgggt 4800
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 4860
taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 4920
tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 4980
tttgatcttt tctacgggggt ctgacgctca gtggaacgaa aactcacgtt aagggatatt 5040
ggctcatgaga ttatcaaaaa ggatcttcac ctagatcctt tttaaattaaa aatgaagttt 5100
taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 5160
tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagttgcct gactcggggg 5220
ggggggggcg tgagggtctgc ctctgtaaga aggtgttgct gactcatacc aggcctgaat 5280
cgccccatca tccagccaga aagtgaggga gccacggttg atgagagctt tggtgtaggt 5340
ggaccagttg gtgattttga acttttgctt tgccacggaa cggctctgcgt tgtcgggaag 5400
atgcgtgac tgatccttca actcagcaaa agttcgattt attcaaaaa gccgccgtcc 5460
cgtcaagtca gcgtaatgct ctgccagtgt tacaaccaat taaccaattc tgattagaaa 5520
aactcatcga gcatcaaatg aaactgcaat ttattcatat caggattatc aataccatat 5580
ttttgaaaaa gccgtttctg taatgaagga gaaaactcac cgaggcagtt ccataggatg 5640
gcaagatcct ggtatcggtc tgcgattccg actcgtccaa catcaataca acctattaat 5700
ttccctcgt caaaaataag gttatcaagt gagaaatcac catgagtgc gactgaatcc 5760
ggtgagaatg gcaaaagctt atgcatttct ttccagactt gttcaacagg ccagccatta 5820
cgctcgtcat caaaatcact cgcatacacc aaaccggtat tcattcgtga ttgcgcctga 5880
gcgagacgaa atacgcgac gctgttaaaa ggacaattac aaacaggaat cgaatgcaac 5940
cggcgcagga aactgccag cgcataca atattttcac ctgaatcagg atattcttct 6000
aatacctgga atgctgtttt cccggggatc gcagtgggtg gtaaccatgc atcatcagga 6060
gtacggataa aatgcttgat ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg 6120
accatctcat ctgtaacatc attggcaacg ctacctttgc catgtttcag aaacaactct 6180
ggcgcacg gcttcccata caatcgatag attgtcgac ctgattgccc gacattatcg 6240
cgagcccatt tatacccata taaatcagca tccatgttgg aatttaatcg cggcctcgag 6300
caagacgttt cccgttgaat atggctcata acacccttg tattactggt tatgtaagca 6360
gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca tcagagattt 6420
tgagacacaa cgtggctttc ccccccccc cattattgaa gcatttatca gggttattgt 6480
ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc 6540
acatttcccc gaaaagtgcc acctgacgct taagaaacca ttattatcat gacattaacc 6600
tataaaaaata ggcgtatcac gaggcccttt cgtc 6634

```

&lt;210&gt; 30

&lt;211&gt; 6538

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 30

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

```

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
 tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
 cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
 agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggcctg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320  
 ggatggggtc ccatttatta ttacaaaatt cacatatata acaacgcctg ccccggtgcc 1380  
 cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500  
 agcggctcat ggtcgtctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgcgc cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tggttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920  
 aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980  
 tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040  
 tgtggaagga ggccaccacc acctgctct gcgccagcga cgccaaggcc tacgacaccg 2100  
 aggtgcacaa cgtgtgggccc acccagcct gcgtgcccac cgaccccaac cccagaggg 2160  
 tgggtgctgg gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220  
 tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280  
 ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340  
 cccaggcctg cccaagggtg agcttcgagc ccatcccaaa ccactactgc gccccgcgcg 2400  
 gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caaggggccc tgcaccaacg 2460  
 tgagcaccgt gcagtgcacc cacggcatcc gcccgtggg gagcaccag ctgctgctga 2520  
 acggtagcct ggccgaggag gaggtggtga tccgcagcgc taacttcgcc gacaacgcca 2580  
 aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgc gctagcgccc 2640  
 actgcaacct gagccgcgcc aagtggaaac acaccctgaa caagatcgtg atcaagctgc 2700  
 gcgagcagtt cggcaacaag accatcgtgt tcaagcacag cagcggcggc gaccccgaga 2760  
 tcgtgaccca cagcttcaac tgcggcggcg agttcttcta ctgcaacagc acccagctgt 2820  
 tcaacagcac ctggttcaac agcacctgga gcaccgagg cagcaacaac accgagggca 2880  
 gcgacaccat caccctgccc tgccgcatca agcagatcat caacatgtgg cagaagggtg 2940  
 gcaaggccat gtacgcccc cccatcagcg gccagatccg ctgcagcagc aacatcaccg 3000  
 gcctgctgct gacccgcgac ggcggaacaa gcaacaacga gagcgagatc ttccgtccgg 3060  
 gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtgaaga 3120  
 tcgagcccct gggcggtggc cccaccaagg ccaagcttac cgtccaggcc cgccagctgc 3180  
 tgagcggcat cgtgcagcag cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc 3240  
 tgctgcagct gaccgtgtgg ggcataaagc agctgcagcc ccgcaccctg gccgtggagc 3300  
 gctacctgaa ggaccagcag cagctcgagc agatctggaa ccacaccacc tggatggagt 3360  
 gggaccgcga gatcaacaac tacaccagcc tgatccacag cctgatcgag gagagccaga 3420  
 accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc agcctgtgga 3480



actgggttcaa catcaccaac tggctgtggt gaggatccag atctgctgtg ccttctagtt 3540  
gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccttgaa ggtgccactc 3600  
ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 3660  
ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 3720  
ggcatgctgg ggatgcggtg ggctctatgg gtaccaggt gctgaagaat tgacctgggt 3780  
cctcctgggc cagaaagaag caggcacatc cccttctctg tgacacaccc tgtccacgcc 3840  
cctgggttctt agttccagcc ccactcatag gacactcata gctcaggagg gctccgcctt 3900  
caatcccacc cgctaaagta cttggagcgg tcttccctc cctcatcagc ccaccaaacc 3960  
aaacctagcc tccaagagt ggaagaaatt aaagcaagat aggctattaa gtgcagaggg 4020  
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat tttaaggcca 4080  
tgattttaagg ccatcatggc cttaatcttc cgcttctctg ctactgact cgctgcgctc 4140  
ggtcggttcgg ctgcggcgag cgggtatcagc tcactcaaag gcggtaatat gggtatccac 4200  
agaatcagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4260  
ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc cgccccctg acgagcatca 4320  
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 4380  
gtttcccccct ggaagctccc tcgtgcgctc tcctgttcg accctgccgc ttaccggata 4440  
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4500  
tctcagttcg gtgtaggteg ttcgctccaa gctgggctgt gtgcacgaac ccccggttca 4560  
gcccagaccg tgcgccttat ccggtacta tcgtcttgag tccaaccgg taagacacga 4620  
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4680  
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4740  
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct ctgtatccgg 4800  
caaacaaacc accgctggta gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag 4860  
aaaaaaagga tctcaagaag atcctttgat ctttctctac gggctctgacg ctcatgtgaa 4920  
cgaaaactca cgttaaggga ttttgggtcat gagattatca aaaaggatct tcacctagat 4980  
cctttttaaat taaaaatgaa gtttttaaat aatctaaagt atatatgagt aaacttgggtc 5040  
tgacagttac caatgcttaa tcagttaggc acctatctca gcgactctgtc tatttctgtt 5100  
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5160  
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5220  
gttgatgaga gctttgttgt aggtggacca gttgggtgat ttgaactttt gctttgccac 5280  
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5340  
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5400  
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5460  
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5520  
tcaccgaggg agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5580  
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc aagttagaaa 5640  
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcac ttctttccag 5700  
acttghtcaa caggccagcc attacgctcg tcatacaaat cactcgcatc aaccaaaccg 5760  
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5820  
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5880  
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgagtg 5940  
gtgagtaacc atgcatcatc aggagtacgg ataaaaatgct tgatgggtcg aagaggcata 6000  
aattccgtca gccagtttag tctgaccatc tcattctgtaa catcattggc aacgctacct 6060  
ttgccatggt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6120  
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6180  
ttggaattta atcgcgccct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6240  
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6300  
tgtgcaatgt aacatcagag attttgagac acaacgtggc ttccccccc ccccatatt 6360  
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6420  
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6480  
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttctgtc 6538

&lt;211&gt; 6541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 31

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagact tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggac acccacgcct gcgtgcccac cgaccccaac cccaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaggcctg cccaagggtg agcttcgagc ccatcccaaa ccactactgc gccccgccc 2400
gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caagggcccc tgcaccaacg 2460
tgagcaccgt gcagtgcacc cacggcatcc gcccgtggt gagcaccag ctgctgctga 2520
acggtagcct ggccgaggag gaggtggtga tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgc cccaacaaca 2640
acacccgcaa gagcatccac atcggccccg gccgcgcctt ctacaccacc ggcgagatca 2700
tcggcgacat ccgccaggcc cactgcaacc tgagccgcgc caagtggaac gacaccctga 2760
acaagatcgt gatcaagctg cgcgagcagt tcggcaacaa gaccatcgtg ttcaagcaca 2820

```

gcagcggcgg cgaccccgag atcgtgaccc acagcttcaa ctgcggcggc gagttcttct 2880  
actgcaacag cgctagcctg cctgcccga tcaagcagat catcaacatg tggcagaagg 2940  
tgggcaaggc catgtacgcc ccccccatca gcgccagat ccgctgcagc agcaacatca 3000  
ccggcctgct gctgaccgcg gacggcggca acagcaacaa cgagagcgag atcttccgtc 3060  
cgggcggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aagggtggtga 3120  
agatcgagcc cctgggcgtg gccccacca aggccaagct taccgtccag gcccgccagc 3180  
tgctgagcgg catcgtgcag cagcagaaca acctgctgcg cgccatcgag gcccgagcagc 3240  
acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgacc ctggccgtgg 3300  
agcgctacct gaaggaccag cagcagctcg agcagatctg gaaccacacc acctggatgg 3360  
agtgggaccg cgagatcaac aactacacca gcctgatcca cagcctgate gaggagagcc 3420  
agaaccagca ggagaagaac gagcaggagc tgctggagct ggacaagtgg gccagcctgt 3480  
ggaactgggt caacatcacc aactggctgt ggtgaggatc cagatctgct gtgccttcta 3540  
gttgccagcc atctgttgtt tgccctccc ccgtgccttc cttgaccctg gaagggtcca 3600  
ctccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg agtaggtgtc 3660  
attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata 3720  
gcaggcatgc tggggatgcg gtgggctcta tgggtaccca ggtgctgaag aattgaccgg 3780  
gttcctcctg ggccagaaag aagcaggcac atccccttct ctgtgacaca ccctgtccac 3840  
gcccctgggt cttagtcca gcccactca taggacactc atagctcagg agggctccgc 3900  
cttcaatccc acccgctaaa gtacttgag cggctctctc ctccctcatc agcccaccaa 3960  
accaaaccta gcctccaaga gtgggaagaa attaaagcaa gataggctat taagtgcaga 4020  
gggagagaaa atgcctccaa catgtgagga agtaatgaga gaaatcatag aattttaagg 4080  
ccatgattta aggccatcat ggccttaatc ttccgcttcc tcgctcactg actcgctgcg 4140  
ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtta tacggttatc 4200  
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 4260  
gaaccgtaaa aaggccgcgt tgctggcgtt ttcccatagg ctccgcccc ctgacgagca 4320  
tcacaaaaat cgacgtcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca 4380  
ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgacctgc cgcttaccgg 4440  
atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct cacgctgtag 4500  
gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 4560  
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaaagaca 4620  
cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 4680  
cgggtgctaca gagttcttga agtgggtggc taactacggc tacactagaa gaacagtatt 4740  
tggtatctgc gctctgctga agccagttac ctbcggaaaa agagtbggta gctcttgatc 4800  
cggcaaaaa accaccgtg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg 4860  
cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg 4920  
gaacgaaaac tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcaccta 4980  
gatectttta aattaaaaat gaagtlttaa atcaatctaa agtatatatg agtaaacttg 5040  
gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 5100  
ttcatccata gttgcctgac tcgggggggg ggggcgctga ggtctgcctc gtgaagaagg 5160  
tggtgctgac tcataccagg cctgaatcgc cccatcatcc agccagaaag tgaggagacc 5220  
acggttgatg agagctttgt ttaggtgga ccagttggtg attttgaact tttgctttgc 5280  
cacggaacgg tctgcgttgt cgggaagatg cgtgatctga tccttcaact cagcaaaagt 5340  
tcgatttatt caacaaagcc gccgtcccgt caagtcagcg taatgctctg ccagtgttac 5400  
aaccaattaa ccaattctga ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta 5460  
ttcatatcag gattatcaat accatathtt tgaaaaagcc gtttctgtaa tgaaggagaa 5520  
aactcaccga ggcagttcca taggatggca agatcctggt atcggtctgc gattccgact 5580  
cgccaacat caatacaacc tattaatttc ccctcgtcaa aaataagggt atcaagttag 5640  
aatcaccat gagtgacgac tgaatccggg gagaatggca aaagcttatg catttcttct 5700  
cagacttggt caacaggcca gccattacgc tcgtcatcaa aatcactcgc atcaaccaa 5760  
ccgttattca ttcgtgattg cgcctgagcg agacgaaata cgcgatcgct gttaaaagga 5820  
caattacaaa caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata 5880  
ttttcacctg aatcaggata ttcttctaata acctggaatg ctgttttccc ggggatcgca 5940  
gtggtgagta accatgcac atcaggagta cggataaaat gcttgatggt cggaagaggc 6000  
ataaatccg tcagccagtt tagtctgacc atctcatctg taacatcatt ggcaacgcta 6060  
cctttgccat gtttcagaaa caactctggc gcategggct tccatacaa tcgatagatt 6120



```

gtcgcacctg attgccccgac attatcgcgga gcccatthtat acccatataa atcagcatcc 6180
atgttggaat ttaatcgcgg cctcgagcaa gacgtttccc gttgaatatg gctcataaca 6240
ccccctgtat tactgttttat gtaagcagac agttttattg ttcattgatga tataattttta 6300
tcttgtgcaa tgtaacatca gagatthtga gacacaacgt ggctttcccc ccccccccat 6360
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6420
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 6480
gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt 6540
c

```

&lt;210&gt; 32

&lt;211&gt; 6577

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 32

```

tcgcgcggtt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcggtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc battaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttgc gagacgccat ccacgtgtgt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtct ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tccctttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100

```

```

aggtgcacaa cgtgtgggccc acccagcct gcgtgcccac cgaccccaac cccagaggagg 2160
tggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaagggtg cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 2460
aggcctgccc caagggtgagc ttcgagccca tccccaacca ctactgcgcc cccgcccggct 2520
tcgccatcct gaagtgcaag gacaagaagt tcaacggcaa gggcccctgc accaacgtga 2580
gcaccgtgca gtgcacccac ggcattccgc ccgtggtgag caccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcgct agcgcctact 2760
gcaacctgag ccgcgccaag tggaaacgaca ccctgaacaa gatcgtgate aagctgcgcg 2820
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcgccgac cccgagatcg 2880
tgaccacag cttcaactgc ggcggcgagt tcttctactg caacagcgt agcctgccct 2940
gccgcatcaa gcagatcatc aacatgtggc agaagggtgg caaggccatg tacgcccccc 3000
ccatcagcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg acccgcgacg 3060
gcggcaacag caacaacgag agcgagatct tccgtccggg cggcgccgac atgcgcgaca 3120
actggcgag cgagctgtac aagtacaagg tgggtgaagat cgagcccctg ggcgtggccc 3180
ccaccaaggc caagcttacc gtccaggccc gccagctgct gagcggcatc gtgcagcagc 3240
agaacaacct gctgcgcgcc atcgaggccc agcagcacct gctgcagctg accgtgtggg 3300
gcatcaagca gctgcaggcc cgcaccctgg ccgtggagcg ctacctgaag gaccagcagc 3360
agctcgagca gatctggaac cacaccacct ggatggagtg ggaccgcgag atcaacaact 3420
acaccagcct gatccacagc ctgatcgagg agagccagaa ccagcaggag aagaacgagc 3480
aggagctgct ggagctggac aagtgggcca gcctgtggaa ctggttcaac atcaccaact 3540
ggctgtggtg aggatccaga tctgctgtgc cttctagtgt ccagccatct gttgtttgcc 3600
cctccccgt gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa 3660
atgaggaaat tgcatcgcat tgtctgagta ggtgtcatc tattctgggg ggtggggtgg 3720
ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg 3780
gctctatggg taccaggtg ctgaagaatt gaccgggttc ctctggggcc agaaagaagc 3840
aggcacatcc ccttctctgt gacacacct gtccacgccc ctggttctta gttccagccc 3900
cactcatagg aactcatag ctgaggagg ctccgccttc aatccacccc gctaaagtac 3960
ttggagcggc ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg 4020
gaagaaatta aagcaagata ggctattaag tgcagagggg gagaaaatgc ctccaacatg 4080
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4140
ttaatcttcc gcttctctgc tcaactgact gctgcgctcg gtcgttcggc tgcggcgagc 4200
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4260
aaagaacatg tgagcaaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct 4320
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4380
gaggtggcga aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4440
cgtgcgtctt cctgttccga ccctgcgct taccggatac ctgtccgcct ttctcccttc 4500
gggaagcgtg gcgctttctc atagctcacg ctgtagggtat ctgagttcgg tgtaggctgt 4560
tcgctccaag ctgggctgtg tgcacgaacc cccggttcag cccgaccgct gcgccttata 4620
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4680
cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg 4740
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4800
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4860
cgggtggttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4920
tcctttgatc ttttctacgg ggtctgacgc tcagtggaaac gaaaactcac gttaagggat 4980
tttgggtcat agattatcaa aaaggatctt cacctagatc ctttttaatt aaaaatgaag 5040
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5100
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5160
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5220
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5280
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5340
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5400

```

```

tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5520
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gtcccatagg 5580
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5640
aatttccctt cgtcaaaaat aagggtatca agtgagaaat caccatgagt gacgactgaa 5700
tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5820
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaacagg aatcgaatgc 5880
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 5940
tctaatacct ggaatgctgt tttcccgagg atcgcagtggt tgagtaacca tgcattcatca 6000
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6060
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6180
tcgcgagccc atttataccc atataaatca gcatccatgt tgggaatttaa tcgcggcctc 6240
gagcaagacg tttcccgctt aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct tcccccccc cccattattt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgat cagcaggccc tttcgtc 6577

```

&lt;210&gt; 33

&lt;211&gt; 6538

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 33

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcggggtg tcgggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgacccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc ggcatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtcctc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggctc atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaate cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatatata acaacgccgt cccccgtgcc 1380

```

cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgcctgg	gctggcgcctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccg	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtgggccc	acccacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tggtgctggt	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcat	gatcatggag	aagggcgaga	tcaagaactg	cagcttcgct	agcaccagct	2400
gcaacaccag	cgtgatcacc	caggcctgcc	ccaaggtgag	cttcgagccc	atccccaacc	2460
actactgcgc	ccccgccggc	ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	2520
agggccccctg	caccaacgtg	agcaccgtgc	agtgcaccca	cggcatccgc	cccggtggtga	2580
gcacccagct	gctgctgaac	ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	2640
acttcgcccga	caacgccaaag	gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	2700
gcacccgcgc	tagcgcgccac	tgcaacctga	gccgcgccaa	gtggaacgac	accctgaaca	2760
agatcgtgat	caagctgcgc	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2820
gcggcgggcga	ccccgagatc	gtgaccacac	gcttcaactg	cggcgggcgag	ttcttctact	2880
gcaacagcgc	tagcctgccc	tgcgcgatca	agcagatcat	caacatgtgg	cagaaggtgg	2940
gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	ctgcagcagc	aacatcaccg	3000
gcctgctgct	gacccgcgac	ggcggcaaca	gcaacaacga	gagcgagatc	ttccgtccgg	3060
gcggcgggcga	catgcgcgac	aactggcgca	gcgagctgta	caagtacaag	gtggtgaaga	3120
tcgagcccct	gggcgtggcc	cccaccaagg	ccaagcttac	cgtccaggcc	cgccagctgc	3180
tgagcggcat	cgtgcagcag	cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	3240
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcaggc	ccgcaccctg	gccgtggagc	3300
gctacctgaa	ggaccagcag	cagctcgagc	agatctggaa	ccacaccacc	tggatggagt	3360
gggaccgcga	gatcaacaac	tacaccagcc	tgatccacag	cctgatcgag	gagagccaga	3420
accagcagga	gaagaacgag	caggagctgc	tggagctgga	caagtgggccc	agcctgtgga	3480
actggttcaa	catcaccaac	tggctgtggt	gaggatccag	atctgctgtg	ccttctagtt	3540
gccagccatc	tgttgtttgc	ccctcccccg	tgccttcctt	gaccctggaa	ggtgccactc	3600
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3660
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3720
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccagagt	gctgaagaat	tgacccggtt	3780
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	3840
cctggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	3900
caatcccacc	cgctaaagta	cttgagcgcg	tctctccctc	cctcatcagc	ccaccaaacc	3960
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	4020
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	4080
tgattttaagg	ccatcatggc	cttaatcttc	cgttctctcg	ctcactgact	cgctgcgctc	4140
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatac	ggttatccac	4200
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4260
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgcgcccttg	acgagcatca	4320
caaaaatcga	cgctcaagtc	agagggtggcg	aaacccgaca	ggactataaa	gataccaggc	4380
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgtttccg	accctgccgc	ttaccggata	4440
cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	4500
tctcagttcg	gtgtaggtcg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	4560
gcccgaaccgc	tgcgccttat	ccggttaacta	tgcgtcttgag	tccaacccgg	taagacacga	4620
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgagggt	atgtaggcgg	4680



```

tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4740
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4800
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 4860
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 4920
cgaaaactca cgttaaggga ttttggatcat gagattatca aaaaggatct tcacctagat 4980
cctttttaa ataaaaatgaa gtttttaaate aatctaaagt atatatgagt aaacttggtc 5040
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgactctgtc tatttcggtc 5100
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5160
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5220
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5280
ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 5340
atthattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5400
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5460
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5520
tcaccgagge agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5580
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc aagtgagaaa 5640
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5700
acttgttcaa caggccagcc attacgctcg tcatacaaat cactcgcac caccaaaaccg 5760
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5820
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatattt 5880
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgagtg 5940
gtgagtaacc atgcatcatc aggagtacgg ataaaaatgct tgatggctcg aagaggcata 6000
aattccgtca gccagtttag tctgaccatc tcactctgtaa catcattggc aacgctacct 6060
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6120
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6180
ttggaattta atcgcgccct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6240
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6300
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 6360
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6420
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6480
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6538

```

&lt;210&gt; 34

&lt;211&gt; 6460

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 34

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatagcgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720

```

catcaatggg	cgtggatagc	ggtttgactc	acgggggattt	ccaagtctcc	acccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtatacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggagggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcgc	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttataaatt	cacatataca	acaacgccgt	ccccgtgcc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgcgtgg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccg	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	accacgcct	gcgtgcccac	cgaccccaac	cccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgacct	2280
ccctgtgcgt	gagcctgaag	tgcaccgacg	ctagcaccag	ctgcaacacc	agcgtgatca	2340
cccaggcctg	ccccaaagg	agcttcgagc	ccatcccaaa	ccactactgc	gcccccgccg	2400
gcttcgccat	cctgaagtgc	aaggacaaga	agttcaacgg	caaggggccc	tgcaccaacg	2460
tgagcacctg	gcagtgacc	cacggcatcc	gccccgtgg	gagcaccag	ctgctgctga	2520
acggtagcct	ggccgaggag	gaggtggtga	tccgcagcgc	taacttcgcc	gacaacgcca	2580
aggtgatcat	cgtgcagctg	aacgagagcg	tggagatcaa	ctgcaccgcg	gctagcgccc	2640
actgcaacct	gagccgcgcc	aagtggaaacg	acaccctgaa	caagatcgtg	atcaagctgc	2700
gcgagcagtt	cggcaacaag	accatcgtgt	tcaagcacag	cagcggcggc	gaccccgaga	2760
tcgtgacca	cagcttcaac	tgcggcgggc	agttcttcta	ctgcaacagc	gctagcctgc	2820
cctgccgcat	caagcagatc	atcaacatgt	ggcagaaggt	gggcaaggcc	atgtacgccc	2880
cccccatcag	cggccagatc	cgtgcagca	gcaacatcac	cggcctgctg	ctgacccgcg	2940
acggcgga	cagcaacaac	gagagcgaga	tcttcggtcc	gggcggcggc	gacatgcgcg	3000
acaactggcg	cagcgagctg	tacaagtaca	aggtgggtgaa	gatcgagccc	ctgggcgtgg	3060
ccccaccaa	ggccaagctt	accgtccagg	cccgcagct	gctgagcggc	atcgtgcagc	3120
agcagaacaa	cctgctgcgc	gccatcgagg	cccagcagca	cctgctgcag	ctgaccgtgt	3180
ggggcatcaa	gcagctgcag	gcccgcaccc	tggccgtgga	gcgctacctg	aaggaccagc	3240
agcagctcga	gcagatctgg	aaccacacca	cctggatgga	gtgggaccgc	gagatcaaca	3300
actacaccag	cctgatccac	agcctgatcg	aggagagcca	gaaccagcag	gagaagaacg	3360
agcaggagct	gctggagctg	gacaagtggg	ccagcctgtg	gaactgggtc	aacatcacca	3420
actggctgtg	gtgaggatcc	agatctgctg	tgccttctag	ttgccagcca	tctgttgttt	3480
gccccctccc	cgtgccttcc	ttgaccctgg	aaggtgccac	tcccactgtc	ctttccta	3540
aaaatgagga	aattgcatcg	cattgtctga	gtaggtgtca	ttctattctg	gggggtgggg	3600
tggggcagga	cagcaagggg	gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	3660
tgggtcttat	gggtacccag	gtgctgaaga	attgaccggg	ttcctcctgg	gccagaaaga	3720
agcaggcaca	tcccccttct	tgtgacacac	cctgtccacg	ccccctgggt	ttagttccag	3780
ccccactcat	aggacactca	tagctcagga	gggtccgcgc	ttcaatccca	cccgttaaag	3840
tacttggagc	ggtctctccc	tccctcatca	gcccacaaa	ccaaacctag	cctccaagag	3900
tgggaagaaa	ttaaagcaag	ataggctatt	aagtgcagag	ggagagaaaa	tgcctccaac	3960
atgtgaggaa	gtaatgagag	aatcataga	attttaaggc	catgatttaa	ggccatcatg	4020

```

gccttaattct tccgcttcct cgctcactga ctcgctgcgc tcggctcggtc ggctgcggcg 4080
agcgggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc 4140
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt 4200
gctggcggtt ttccataggt tccgcccccc tgacgagcat cacaaaaatc gacgctcaag 4260
tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc 4320
cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc 4380
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt 4440
cgttcgctcc aagctgggct gtgtgcacga acccccgcgt cagcccgacc gctgcgcctt 4500
atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc 4560
agccactggg aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa 4620
gtggtggcct aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa 4680
gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg 4740
tagcgggtgg ttttttggtt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga 4800
agatcctttg atcttttcta cggggtctga cgctcagtg aacgaaaact cacgttaagg 4860
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg 4920
aagtttttaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt 4980
aatcagtgag gcacctatct cagcgatctg tctatctcgt tcatccatag ttgcctgact 5040
cggggggggg gggcgctgag gtctgcctcg tgaagaagg gtgtgctgact cataccaggc 5100
ctgaatcgcc ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt 5160
gtaggtggac cagttggtga ttttgaactt ttgctttgcc acggaacgg ctgcgttgtc 5220
gggaagatgc gtgatctgat ctttcaactc agcaaaagt cgatttatc aacaaagccg 5280
ccgtcccgtc aagtcagcgt aatgctctgc cagtgttaca accaatata caattctgat 5340
tagaaaaact catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata 5400
ccatattttt gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat 5460
aggatggcaa gatcctggta tcgggtctgcg attccgactc gtccaacatc aatacaacct 5520
attaatttcc cctcgtcaaa aataagggtt tcaagtgaga aatcaccatg agtgacgact 5580
gaatccggtg agaatggcaa agccttatgc atttctttcc agacttggtc aacaggccag 5640
ccattacgct cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tctgattgct 5700
gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa 5760
tgcaaccggc gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat 5820
tcttctaata cctggaatgc tgttttcccc gggatcgagc tggtagagta ccatgcatca 5880
tcaggagtac ggataaaatg cttgatggtc ggaagaggca taaattccgt cagccagttt 5940
agtctgacca tctcatctgt aacatcattg gcaacgctac ctttgccatg tttcagaaac 6000
aactctggcg catcgggctt cccatacaat cgatagattg tgcacactga ttgcccgaac 6060
ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc 6120
ctcgagcaag acgtttcccc ttgaatatgg ctcataacac cccttgattt actgtttatg 6180
taagcagaca gttttattgt tcatgatgat atatttttat cttgtgcaat gtaacatcag 6240
agattttgag acacaacgtg gctttcccc cccccccatt attgaagcat ttatcagggg 6300
tattgtctca tgagcggata cataattgaa tgtatttaga aaaataaaca aataggggtt 6360
ccgcgcacat tccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 6420
ttaacctata aaaataggcg tatcacgagg ccctttcgtc 6460

```

&lt;210&gt; 35

&lt;211&gt; 7987

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 35

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcggtggg 120

```



acggggcggg	tgacgtagta	gtgtggcgga	agtgtgatgt	tgcaagtgtg	gcggaacaca	180
tgtaagcgac	ggatgtggca	aaagtgacgt	ttttgggtgtg	cgccgggtgta	cacaggaagt	240
gacaattttc	gcgcgggtttt	aggcggtatgt	tgtagtaaat	ttgggcgtaa	ccgagtaaga	300
tttggccatt	ttcgcgggaa	aactgaataa	gaggaagtga	aatctgaata	attttgtgtt	360
actcatagcg	cgtaatatatt	gtctagggcc	gcggggactt	tgaccgttta	cgtggagact	420
cgcccagggtg	tttttctcag	gtgttttccg	cgttccgggt	caaagtggc	gttttattat	480
tatagtcagt	acgtaccagt	gcactggcct	agagcgggcc	cattgcatac	gttgtatcca	540
tatcataata	tgtacattta	tattggctca	tgtccaacat	taccgccatg	ttgacattga	600
ttattgacta	gttattaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	660
gagttccgcg	ttacataact	tacggtaa	ggcccgccctg	gctgaccgcc	caacgacccc	720
cgcccattga	cgtcaataat	gacgtatgtt	cccatagtaa	cgccaatagg	gactttccat	780
tgacgtcaat	gggtggagta	tttacggtaa	actgccact	tggcagtaca	tcaagtgtat	840
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggccgc	ctggcattat	900
gcccagtaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcac	960
gctattacca	tggatgatcg	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	1020
tcacggggat	ttccaagtct	ccacccatt	gacgtcaatg	ggagtttgtt	ttggcaccaa	1080
aatcaacggg	actttccaaa	atgtcgtaac	aactccgcc	cattgacgca	aatggcggt	1140
aggcgtgtac	gggtggagggt	ctatataagc	agagctcgtt	tagtgaaccg	tcagatcgcc	1200
tggagacgcc	atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	1260
cgtcacgctc	gtcgacacgt	gtgatcagat	atcgcgggccg	ctctagaatg	cgcgtgaagg	1320
agaagtacca	gcacctgtgg	cgctggggct	ggcgtgggg	caccatgctg	ctgggcatgc	1380
tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgcacaacgt	gtgggcccacc	cacgcctgcg	tgcccaccga	ccccaacccc	caggagggtg	1560
tgctgggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatgggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgctgaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacgcta	gcaagaactg	cagcttcaac	atcagcacca	1740
gcacccgcg	caaggtgcag	aaggagtacg	ccttcttcta	caagctggac	atcatcccca	1800
tcgacaacga	caccaccagc	tacagcctga	ccagctgcaa	caccagcgtg	atcaccagg	1860
cctgccccaa	ggtgagcttc	gagcccatcc	ccaaccacta	ctgcgcccc	gccggcttcg	1920
ccatcctgaa	gtgcaaggac	aagaagttca	acggcaagg	ccctgcacc	aacgtgagca	1980
ccgtgcagtg	caccacggc	atccgccccg	tgggtgagcac	ccagctgctg	ctgaacggta	2040
gcctggccga	ggaggagggtg	gtgatccgca	gcgctaactt	cgccgacaac	gccaagggtga	2100
tcacgtgca	gctgaacgag	agcgtggaga	tcaactgcac	ccgccccaac	aacaacaccc	2160
gcaagagcat	ccacatcggc	cccggccgcg	ccttctacac	caccggcgag	atcatcggcg	2220
acatccgcca	ggcccactgc	aacctgagcc	gcgccaagt	gaacgacacc	ctgaacaaga	2280
tcgtgatcaa	gctgcgcgag	cagttcggca	acaagaccat	cgtgttcaag	cacagcagcg	2340
gcggcgaccc	cgagatcgtg	acccacagct	tcaactgcgg	cggcgagttc	ttctactgca	2400
acagcaccca	gctgttcaac	agcacctgggt	tcaacagcac	ctggagcacc	gagggcagca	2460
acaacaccga	gggcagcgac	accatcacc	tgccctgccc	catcaagcag	atcatcaaca	2520
tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcgccag	atccgctgca	2580
gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	aacgagagcg	2640
agatcttccg	tccgggccc	ggcgacatgc	gcgacaactg	gogcagcgag	ctgtacaagt	2700
acaagggtggt	gaagatcgag	cccctgggcg	tggcccccac	caaggccaag	cttaccgtcc	2760
aggcccgcga	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	cgcgccatcg	2820
aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	caggcccgcga	2880
ccctggccgt	ggagcgctac	ctgaaggacc	agcagcagct	cgagcagatc	tgggaaccaca	2940
ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgate	cacagcctga	3000
tcgaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	ctggacaagt	3060
gggcccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtgagga	tccagatctg	3120
ctgtgccttc	tagttgccag	ccatctgttg	tttgccctc	ccccgtgcct	tccttgaccc	3180
tggaagggtgc	cactcccact	gtcctttcct	aataaaatga	ggaaattgca	tcgcattgtc	3240
tgagtaggtg	tcattctatt	ctgggggggtg	gggtggggca	gcacagcaag	ggggaggatt	3300
gggaagacaa	tagcaggcat	gctggggatg	cggtgggctc	tatgggtacc	cagggccgca	3360
taacttcgta	taatgtatgc	tatacgaagt	tataagatct	gtactgaaat	gtgtgggcgt	3420

ggcttaaggg tgggaaagaa tatataaggt gggggtctta tgtagttttg tatctgtttt 3480  
gcagcagccg ccgcccgcac gagcaccaac tcgtttgatg gaagcattgt gagctcatat 3540  
ttgacaacgc gcatgcccc atggggccggg gtgcgtcaga atgtgatggg ctccagcatt 3600  
gatggtcgcc ccgtcctgcc cgcaaactct actacctga cctacgagac cgtgtctgga 3660  
acgccgttgg agactgcagc ctccgccgcc gcttcagccg ctgcagccac cgcccgcggg 3720  
attgtgactg actttgcttt cctgagcccg cttgcaagca gtgcagcttc ccgttcatec 3780  
gcccgcgatg acaagttgac ggctcttttg gcacaattgg attctttgac ccgggaactt 3840  
aatgtcgttt ctacagcagc gttggatctg cgccagcagg tttctgccct gaaggcttcc 3900  
tcccccccca atgcgggttta aaacataaat aaaaaaccag actctgtttg gatttggtatc 3960  
aagcaagtgt cttgctgtct ttatttaggg gttttgcgcg cgcggtaggc ccgggaccag 4020  
cgggtctcggg cgttgagggg cctgtgtatt ttttccagga cgtggtaaaag gtgactctgg 4080  
atgttcagat acatgggcat aagcccgctc ctgggggtgga ggtagacca ctgcagagct 4140  
tcatgctgcg ggggtggtgt gtagatgatc cagtcgtagc aggagcgtg ggctggtgc 4200  
ctaaaaatgt ctttcagtag caagctgatt gccaggggca ggcccttggt gtaagtgttt 4260  
acaaagcggg taagctggga tgggtgcata cgtggggata tgagatgcat cttggactgt 4320  
attttttaggt tggctatgtt ccagccata tccctccggg gattcatgtt gtgcagaacc 4380  
accagcacag tgtatccggg gcacttgga aatttgcctat gtagcttaga aggaaatgcg 4440  
tggaagaact tggagacgcc cttgtgacct ccaagatttt ccatgcattc gtccataatg 4500  
atggcaatgg gccacgggc ggccgctgg gcgaagatat ttctgggatc actaacgtca 4560  
tagttgtgtt ccaggatgag atcgtcatag gccattttta caaagcgcgg gcggagggtg 4620  
ccagactgcg gtataatggg tccatccggc ccagggcggt agttaccctc acagatttgc 4680  
atttcccacg ctttgagtgc agatgggggg atcatgtcta cctgcggggc gatgaagaaa 4740  
acggtttccg gggtagggga gatcagctgg gaagaaagca ggttcctgag cagctgcgac 4800  
ttaccgcagc cggtagggcc gtaaatcaca cctattaccg gctgcaactg gtagttaaga 4860  
gagctgcagc tgccgtcatc cctgagcagg ggggccactt cgttaagcat gtccctgact 4920  
cgcattgttt cctgaccaaa atccgccaga aggcgtctgc cgccagcga tagcagttct 4980  
tgcaaggaag caaagttttt caacggtttg agaccgtccg ccgtaggcat gcttttgagc 5040  
gtttgaccaa gcagttccag gcgggtccac agctcgggca cctgctctac ggcatctcga 5100  
tccagcatat ctctcgttt cgcggttgg ggcggtttc gctgtacggc agtagtcggg 5160  
gctcgtccag acgggccagg gtcattgtct tccacggggc cagggtcctc gtcagcgtag 5220  
tctgggtcac ggtgaagggg tgcgtccgg gctgcgcgt ggccaggggt cgcttgaggc 5280  
tgggtcctgct ggtgctgaag cgtgcgggt cttcgccctg cgcgtcggcc aggtagcatt 5340  
tgaccatggt gtcatagtcc agccctccg cgcggtggcc cttggcgcg agcttgccct 5400  
tggaggaggc gccgcacgag gggcagtgca gacttttgag ggcgtagagc ttgggcgcga 5460  
gaaataccga ttccggggag taggcattcc cgccgcaggc ccgcagacg gtctcgcatt 5520  
ccacgagcca ggtgagctct ggccgttcgg ggtcaaaaac caggtttccc ccatgctttt 5580  
tgatgcgttt cttacctctg gtttccatga gccgggtgtc acgctcgggt acgaaaaggc 5640  
tgtccgtgtc cccgtataca gacttgagag gcctgtcctc gaggcgtgtt ccgcgggtcct 5700  
cctcgtatag aaactcggac cactctgaga caaaggctcg cgtccaggcc agcacgaagg 5760  
aggctaagtg ggaggggtag cggtcgttgt ccactagggg gtccactcgc tccagggtgt 5820  
gaagacacat gtcgccctct tcggcatcaa ggaagggtgat tggttttagt gtgtaggcca 5880  
cgtgaccggg tgttccctgaa ggggggctat aaaagggggg gggggcgctg tcgtcctcac 5940  
tctcttccgc atcgtgtgtt gcgagggcca gctgttgggg tgagtcgacg cgaggctgga 6000  
tggccttccc cattatgatt cttctcgtct ccggcgccat cgggatgccc gcgttgagg 6060  
ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaaggc cagcaaaagg 6120  
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg 6180  
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa ccgcagagga ctataaagat 6240  
accaggcgtt tccccctgga agctccctcg tgcgtctcc tgttccgacc ctgcccgtta 6300  
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 6360  
gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc 6420  
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccgggtaa 6480  
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggatg 6540  
taggcgggtg tacagagtcc ttgaagtggg ggcctaacta cggctacact agaaggacag 6600  
tatttggtat ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 6660  
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 6720

```

cgcgagaaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 6780
agtggaaacga aaactcacgt taagggatct tgggtcatgag attatcaaaa aggatcttca 6840
cctagatcct tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa 6900
cttgggtctga cagttaccaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat 6960
ttcgttcate catagttgcc tgactccccg tcgtgtagat aactacgata cgggaggggt 7020
taccatctgg cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 7080
tatcagcaat aaaccagcca gccggaaggg ccgagcgag aagtggctct gcaactttat 7140
ccgcctccat ccagttctatt aattgttgcc gggaaagctag agtaagtagt tcgccagtta 7200
atagtttgcg caacgttggt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 7260
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 7320
tgtgcaaaaa agcgggttagc tccttcgggtc ctccgatcgt tgtcagaagt aagttggccg 7380
cagtgttatc actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg 7440
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 7500
ggcgaccgag ttgctcttgc ccggcgtaaa cacgggataa taccgcgcca catagcagaa 7560
ctttaaaagt gctcatcatt ggaaaacggt cttcggggcg aaaactctca aggatcttac 7620
cgctgttgag atccagttcg atgtaaccca ctctgtcacc caactgatct tcagcatctt 7680
ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 7740
gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 7800
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 7860
aacaatatag ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 7920
ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtcttcaag 7980
aattggtt                                     7987

```

&lt;210&gt; 36

&lt;211&gt; 7948

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 36

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttggt acgtggcgcg gggcggtgga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttgggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttggtgt 360
actcatagcg cgtaatatct gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcaactggcc agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatggt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gccagtagca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcate 960
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttggt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggg 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260

```



```

cgtcaccgtc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgetggggct ggcgtggggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgccacc gagaagctgt gggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtggggccacc cacgcctgcg tgcccaccga ccccaacccc caggaggtgg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcbtgaag ctgaccccc 1680
tgtgcbtgaag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740
gccgcatgat catggagaag ggcgagatca agaactgcag ctctgctagc accagctgca 1800
acaccagcgt gatcaccag gcctgcccc aaggtagctt cgagcccatc cccaaccact 1860
actgcgcccc cgccggcttc gccatcctga agtgcaagga caagaagtcc aacggcaagg 1920
gccccctgcac caacgtgagc accgtgcagt gcaccacagg catccgcccc gtggtgagca 1980
cccagctgct gctgaacggc agcctggccg aggaggagg ggtgatccgc agcgctaact 2040
tcgcccacaa cgccaagggt atcatcgtgc agctgaacga gagcgtggag atcaactgca 2100
cccgcccaa caacaacacc cgcaagagca tccacatcgg ccccgccgc gccttctaca 2160
ccaccggcga gatcatcggc gacatccgcc agggccactg caacctgagc cgcgccaagt 2220
ggaacgacac cctgaacaag atcgtgatca agctgcgcga gcagttcggc aacaagacca 2280
tcgtgttcaa gcacagcagc ggcggcgacc ccgagatcgt gaccacagc ttcaactgca 2340
gcggcgagtt cttctactgc aacagcacc agctgttcaa cagcacctgg ttcaacagca 2400
cctggagcac cgagggcagc aacaacacc agggcagcga caccatcacc ctgccctgcc 2460
gcatacaagca gatcatcaac atgtggcaga aggtgggcaa ggccatgtac gccccccca 2520
tcagcggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc cgcgacggcg 2580
gcaacagcaa caacgagagc gagatcttcc gtccggcgcg cggcgacatg cgcgacaact 2640
ggcgagcga gctgtacaag tacaagggtg tgaagatcga gcccctgggc gtggcccca 2700
ccaaggccaa gcttaccgtc caggcccgcc agctgctgag cggcatcgtg cagcagcaga 2760
acaacctgct gcgcgccatc gagggccagc agcacctgct gcagctgacc gtgtggggca 2820
tcaagcagct gcaggcccg ccctggccg tggagcgcta cctgaaggac cagcagcagc 2880
tcgagcagat ctggaaccac accacctgga tggagtggga ccgcgagatc aacaactaca 2940
ccagcctgat ccacagcctg atcgaggaga gccagaacca gcaggagaag aacgagcagg 3000
agctgctgga gctggacaag tgggcccagc tgtggaactg gttcaacatc accaactggc 3060
tgtggtgagg atccagatct gctgtgcctt ctagttgcca gccatctgtt gtttgccct 3120
ccccctgcc ttctctgacc ctggaagggt ccactcccac tgcctttcc taataaaatg 3180
aggaaattgc atcgcatgtt ctgagtaggt gtcattctat tctggggggg ggggtggggc 3240
agcacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct 3300
ctatgggtac ccagggccgc ataacttcgt ataattgatg ctatacgaag ttataagatc 3360
tgtactgaaa tgtgtggggc tggcttaagg gtgggaaaga atatataagg tgggggtctt 3420
atgtagtttt gtatctgttt tgcagcagcc gccgcgcca tgagcaccaa ctctgtttgat 3480
ggaagcattg tgagctcata tttgacaacg cgcagcccc catggggcgg ggtgcgtcag 3540
aatgtgatgg gctccagcat tgatggtcgc ccgctcctgc ccgcaaactc tactaccttg 3600
acctacgaga ccgtgtctgg aacgcgctt gagactgcag cctccgccc cgttccagcc 3660
gctgcagcca ccgcccgcgg gattgtgact gactttgctt tccctgagccc gcttgcaagc 3720
agtgcagctt cccgttcatc cgcccgcgat gacaagtga cggctctttt ggcacaattg 3780
gattctttga cccgggaact taatgtcgtt tctcagcagc tgttggtatc gcgccagcag 3840
gtttctgccc tgaaggcttc ctccccctcc aatgcggttt aaaacataaa taaaaaacca 3900
gactctgttt ggatttggat caagcaagtg tcttgctgct tttatltagg ggttttgcgc 3960
gcgcggtagg cccgggacca gcggtctcgg tcgttgaggg tccctgtgat tttttccagg 4020
acgtggtaaa ggtgactctg gatgttcaga tacatgggca taagcccgtc tctgggggtg 4080
aggtagcacc actgcagagc ttcagctgct ggggtggtgt tgtagatgat ccagtcgtag 4140
caggagcgct gggcgtggtg cctaaaaatg tctttcagta gcaagctgat tgccaggggc 4200
aggcccttgg tgtaagtgtt tacaagcgg ttaagctggg atgggtgcat acgtggggat 4260
atgagatgca tcttggaactg tttttttagg ttggctatgt tcccagccat atccctccgg 4320
ggattcatgt tgtgcagaac caccagcaca gtgtatccgg tgcacttggg aaatttgtca 4380
tgtagcttag aaggaaatgc gtggaagaac ttggagacgc ccttgtgacc tccaagattt 4440
tccatgcatt cgtccataat gatggcaatg ggcacacgg cggcggcctg ggcgaagata 4500
tttctgggat cactaacgtc atagtgtgtg tccaggatga gatcgtcata ggccattttt 4560

```

```

acaaagcgcg ggcggagggt gccagactgc ggtataatgg ttccatccgg cccagggggcg 4620
tagttaccct cacagatttg catttcccac gctttgagtt cagatggggg gatcatgtct 4680
acctgcgggg cgatgaagaa aacggtttcc ggggtagggg agatcagctg ggaagaaagc 4740
aggttcctga gcagctgcga cttaccgcag ccggtggggc cgtaaatac acctattacc 4800
ggctgcaact ggtagttaag agagctgcag ctgccgtcat ccctgagcag gggggccact 4860
tcgttaagca tgtccctgac tcgcatgttt tccctgacca aatccgccag aaggcgctcg 4920
ccgcccagcg atagcagttc ttgcaaggaa gcaaagtttt tcaacgggtt gagaccgtcc 4980
gccgtaggca tgcttttgag cgtttgacca agcagttcca ggcggtccca cagctcggtc 5040
acctgctcta cggcatctcg atccagcata tctcctcggt tcgcggggtg gggcggtttt 5100
cgctgtacgg cagtagtcgg tgctcgteca gacgggccag ggtcatgtct ttccacgggc 5160
gcagggtcct cgtcagcgta gtctgggtca cgggtgaagg gtgcgctccg ggctgcgcgc 5220
tggccagggt gcgcttgagg ctggctcctgc tgggtgctgaa gcgctgccgg tcttcgccct 5280
gcgcgtcggc caggtagcat ttgaccatgg tgtcatagtc cagccccctc gcggcggtggc 5340
ccttggcgcg cagcttgccc ttggaggagg cggcgcacga ggggcagtg c agacttttga 5400
gggcgtagag cttgggcgcg agaaataccg attccgggga gtaggcattc gcgccgcagg 5460
ccccgcagac ggtctcgcat tccacgagcc aggtgagctc tggccgttcg ggggtcaaaaa 5520
ccaggtttcc cccatgcttt ttgatgcgtt tcttacctct ggtttccatg agccgggtgc 5580
cacgctcggt gacgaaaagg ctgtccgtgt ccccgatac agacttgaga ggctgtcct 5640
cgagcggtgt tccgcggtcc tctcgtata gaaactcgga ccactctgag acaaaggctc 5700
gcgtccaggc cagcacgaag gaggctaagt gggaggggta gcggtcgttg tccactaggg 5760
ggtccactcg ctccagggtg tgaagacaca tgtcgccctc ttccggcatca aggaagggtga 5820
ttggttttga ggtgtaggcc acgtgaccgg gtgttcctga aggggggcta taaaaggggg 5880
tggggggcgcg ttctgctcctca ctctcttccg catcgctgtc tgcgagggcc agctgttggg 5940
gtgagtcgac gcgaggctgg atggccttcc ccattatgat tcttctcgct tccggcgga 6000
tcgggatgcc cgcgttgca gccatgctgt ccaggcaggt agatgacgac catcagggac 6060
agcttcaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 6120
ataggctccg cccctctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 6180
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc 6240
ctgttccgac cctgcgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 6300
cgctttctca tagctcacgc tgtaggatc tcagttcggg gtaggtcgtt cgtccaagc 6360
tgggctgtgt gcacgaacce cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 6420
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca 6480
ggattagcag agcgaggat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact 6540
acggctacac tagaaggaca gtatttggt tctgcgctct gctgaagcca gttaccttcg 6600
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt 6660
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 6720
tttctacggg gtctgacgct cagtggaaacg aaactcacg ttaagggatt ttgggtcatga 6780
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaataca 6840
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 6900
ctatctcage gatctgtcta ttctgttcct ccatagttgc ctgactcccc gtcgtgtaga 6960
taactacgat acgggagggc ttaccatctg gccccagtg tgcaatgata ccgcgagacc 7020
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 7080
gaagtggctc tgcaacttta tccgcctcca tccagctctat taattgttg cgggaagcta 7140
gagtaagtag ttcgccagtt aatagtttgc gcaacggtgt tgccattgct gcaggcatcg 7200
tgggtgcacg ctctgcgttt ggtatggctt cattcagctc cggbtcccaa cgatcaaggc 7260
gagttacatg atcccccatg ttgtgcaaaa aagcgggttag ctcttcgggt cctccgatcg 7320
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 7380
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 7440
cattctgaga atagtgtatg cggcgaccga gttgtctctt cccggcgctc acacgggata 7500
ataccgcgcc acatagcaga actttaaaag tgcctcatcat tggaaaacgt tcttcggggc 7560
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacce actcgtgcac 7620
ccaactgate ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 7680
ggcaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 7740
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7800
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7860

```

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7920  
cgaggccctt tcgtcttcaa gaattggt 7948

<210> 37

<211> 7963

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pAdApt containing HIV genes

<400> 37

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60  
ggattgaagc caatatgata atgaggggggt ggagtttgtg acgtggcgcg gggcgtggga 120  
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180  
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240  
gacaattttc gcgcggtttt aggcggatgt ttagtaaat ttgggcgtaa ccgagtaaga 300  
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360  
actcatagcg cgtaatatatt gtctagggcc gcggggactt tgaccgttta cgtggagact 420  
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagtggc gttttattat 480  
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540  
tatcataata tgtacattta tattggctca tgcacaacat taccgccatg ttgacattga 600  
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660  
gagttccgcg ttacataact tacggtaaatt ggcccgcctg gctgaccgcc caacgacccc 720  
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780  
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840  
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900  
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcac 960  
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020  
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080  
aatcaacggg actttccaaa atgtcgtaac aactccgcc cattgacgca aatgggcggt 1140  
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200  
tgagagacgc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260  
cgtcaccgtc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320  
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380  
tgatgatctg cagcgccacc gagaagctgt gggtgaccgt gtactacggc gtgcccgtgt 1440  
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500  
tgacacaacgt gtggggccacc cagcctgcg tgcaccga ccccaacccc caggaggtgg 1560  
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620  
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctgaag ctgaccccc 1680  
tgtgctgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740  
gccgcatgat catggagaag ggcgagatca agaactgcag cttcaacatc agcaccagca 1800  
tccgcggcaa ggtgcagaag gactacgct tcttctacaa gctggacatc atccccatcg 1860  
acaacgacac caccagctac agcctgacca gctgcaacac cagcgtgatc acccaggcct 1920  
gccccagggt gagcttcgag cccatcccca accactactg cgccccgcg ggcttcgcca 1980  
tcctgaagtg caaggacaag aagtccaacg gcaagggccc ctgcaccaac gtgagcaccg 2040  
tgacgtgcac ccacggcatc cgccccgtgg tgagcaccga gctgctgctg aacggtagcc 2100  
tggccgagga ggaggtggtg atccgcagcg ctaacttcgc cgacaacgcc aaggtgatca 2160  
tcgtgcagct gaacgagagc gtggagatca actgcaccgc cgctagcgcc cactgcaacc 2220  
tgagccgcgc caagtgaac gacacctga acaagatcgt gatcaagctg cgcgagcagt 2280  
tcggcaacaa gaccatcgtg ttcaagcaca gcagcggcgg cgaccccgag atcgtgaccc 2340  
acagcttcaa ctgcggcggc gagttcttct actgcaacag caccagctg ttcaacagca 2400  
cctggttcaa cagcacctgg agcaccgagg gcagcaacaa caccgagggc agcgacacca 2460



tcaccctgcc	ctgccgcac	aagcagatca	tcaacatgtg	gcagaagggtg	ggcaaggcca	2520
tgtacgcccc	ccccatcagc	ggccagatcc	gctgcagcag	caacatcacc	ggcctgctgc	2580
tgacccgcga	cggcggcaac	agcaacaacg	agagcgagat	cttccgtccg	ggcggcggcg	2640
acatgcgcga	caactggcgc	agcgagctgt	acaagtacaa	ggtggtgaag	atcgagcccc	2700
tgggcgtggc	ccccaccaag	gccaagctta	ccgtccagge	ccgccagctg	ctgagcggca	2760
tcgtgcagca	gcagaacaac	ctgctgcgcg	ccatcgagge	ccagcagcac	ctgctgcagc	2820
tgaccgtgtg	gggcatcaag	cagctgcagg	cccgcaccct	ggccgtggag	cgctacctga	2880
aggaccagca	gcagctcgag	cagatctgga	accacaccac	ctggatggag	tgggaccgcg	2940
agatcaacaa	ctacaccagc	ctgatccaca	gcctgatcga	ggagagccag	aaccagcagg	3000
agaagaacga	gcaggagctg	ctggagctgg	acaagtgggc	cagcctgtgg	aactggttca	3060
acatcaccaa	ctggctgtgg	tgaggatcca	gatctgctgt	gccttctagt	tgccagccat	3120
ctgttggttg	cccccccccc	gtgccttctt	tgaccctgga	aggtgccact	cccactgtcc	3180
tttccctaata	aaatgaggaa	attgcatcgc	attgtctgag	taggtgtcat	tctattcttg	3240
ggggtggggg	ggggcagcac	agcaaggggg	aggattggga	agacaatagc	aggcatgctg	3300
gggatgcggg	gggctctatg	ggtacccagg	gccgcataac	ttcgtataat	gtatgctata	3360
cgaagtata	agatctgtac	tgaatgtgt	gggcgtggct	taagggtggg	aaagaatata	3420
taagggtggg	gtcttatgta	gttttgtatc	tgttttgcag	cagccgcgcg	cgccatgagc	3480
accaactcgt	ttgatggaag	cattgtgagc	tcatatttga	caacgcgcac	gcccccatgg	3540
gccgggggtg	gtcagaatgt	gatgggctcc	agcattgatg	gtcgccccgt	cctgccccga	3600
aactctacta	ccttgacctc	cgagaccgtg	tctggaacgc	cgttggagac	tgcagcctcc	3660
gccgccgctt	cagccgctgc	agccaccgce	cgccgggattg	tgactgactt	tgctttccctg	3720
agcccgcctg	caagcagtg	agcttcccgt	tcctccgccc	gcgatgacaa	gttgacggct	3780
cttttgccac	aattggattc	tttgaccggg	gaacttaatg	tcgtttctca	gcagctgttg	3840
gatctgcgcc	agcaggtttc	tgccctgaag	gcttccctcc	ctcccaatgc	ggtttaaaac	3900
ataaataaaa	aaccagactc	tgtttggatt	tggatcaagc	aagtgtcttg	ctgtctttat	3960
ttaggggttt	tgcgcgcgcg	gtaggcccgg	gaccagcggg	ctcggtcgtt	gagggctcctg	4020
tgtatttttt	ccaggacgtg	gtaaagggtga	ctctggatgt	tcagatacat	gggcataagc	4080
ccgtctctgg	ggtggaggta	gcaccactgc	agagcttcat	gctgcggggg	ggtgtttag	4140
atgatccagt	cgtagcagga	gcgctggggc	tgggtgcctaa	aatgtcttt	cagtagcaag	4200
ctgattgcca	ggggcaggcc	cttgggtgtaa	gtgtttacaa	agcgggttaag	ctgggatggg	4260
tgcatacgtg	gggatatgag	atgcactctg	gactgtattt	ttaggttggc	tatgttccca	4320
gccatatccc	tccggggatt	catgttgtgc	agaaccacca	gcacagtgtg	tccgggtgcac	4380
ttgggaaatt	tgtcatgtag	cttagaagga	aatgcgtgga	agaacttggg	gacgcccttg	4440
tgacctccaa	gattttccat	gcattcgtcc	ataatgatgg	caatggggcc	acgggcggcg	4500
gcctggggga	agatatttct	gggatcacta	acgtcatagt	tgtgttccag	gatgagatcg	4560
tcataggcca	tttttacaaa	gcgcggggcg	agggtgccag	actgcggtat	aatggttcca	4620
tccggcccag	gggcgtagtt	accctcacag	atttgcattt	cccacgcttt	gagttcagat	4680
ggggggatca	tgtctacctg	cggggcgatg	aagaaaacgg	tttccggggg	aggggagatc	4740
agctgggaag	aaagcagggt	cctgagcagc	tgcgacttac	cgcagccggg	gggcccgtaa	4800
atcacacctc	ttaccggctg	caactggtag	ttaagagagc	tgcagctgcc	gtcatccctg	4860
agcagggggg	ccacttcgtt	aagcatgtcc	ctgactcgca	tgttttccct	gaccaaattc	4920
gccagaaggc	gctcgccgcc	cagcgatagc	agtctctgca	aggaagcaaa	gtttttcaac	4980
ggtttgagac	cgtccgcctg	aggcatgctt	ttgagcgttt	gaccaagcag	ttccaggcgg	5040
tcccacagct	cgggtcacctg	ctctacggca	tctcgatcca	gcatactctc	tctgttccgcg	5100
ggttggggcg	gctttcgtctg	tacggcagta	gtcgggtgctc	gtccagacgg	gccagggtca	5160
tgtctttcca	cgggcgcagg	gtcctcgtca	gcgtagtctg	ggtcacgggtg	aaggggtgcg	5220
ctccgggctg	cgcgctggcc	agggtgcgct	tgaggctggg	cctgctgggtg	ctgaagcgct	5280
gccggtcttc	gcccgtgcgcg	tcggccagggt	agcatttgac	catgggtgtca	tagtccagcc	5340
cctccgcggc	gtggcccttg	gcgcgcagct	tgcccttgga	ggaggcgccg	cacgaggggc	5400
agtgcagact	tttgaggggcg	tagagcttgg	gcgcgagaaa	taccgattcc	ggggagtagg	5460
catccgcgcc	gcaggccccg	cagacggctc	cgcattccac	gagccagggtg	agctctggcc	5520
gttcgggggc	aaaaaccagg	tttcccccat	gctttttgat	gcgtttctta	cctctgggtt	5580
ccatgagccg	gtgtccacgc	tcgggtgacga	aaaggctgtc	cgtgtccccg	tatacagact	5640
tgagaggcct	gtcctcgagc	ggtgttccgc	ggtcctcctc	gtatagaaac	tccgaccact	5700
ctgagacaaa	ggctcgcgtc	caggccagca	cgaaggaggc	taagtgggag	gggtagcggg	5760



```

cgttgtccac taggggggtcc actcgctcca ggggtgtgaag acacatgtcg ccctcttcgg 5820
catcaaggaa ggtgattggt ttgtaggtgt aggccacgtg accgggtggt cctgaagggg 5880
ggctataaaa ggggggtgggg gcgcgttcgt cctcactctc ttccgcatcg ctgtctgcca 5940
gggccagctg ttgggggtgag tcgacgcgag gctggatggc ctccccatt atgattcttc 6000
tcgcttccgg cggcatcggg atgcccgcgt tgcaggccat gctgtccagg caggtagatg 6060
acgaccatca gggacagctt caaggccagc aaaaggccag gaaccgtaaa aaggccgcgt 6120
tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa 6180
gtcagagggt gcgaaaccgg acaggactat aaagatacca ggcgtttccc cctggaagct 6240
ccctcgtgcg ctctctgtgt ccgacctgc cgcttaccgg atacctgtcc gcctttctcc 6300
cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcgggtgtagg 6360
tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct 6420
tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag 6480
cagccactgg taacaggatt agcagagcga ggtatgtagg cgggtgctaca gaggttctga 6540
agtgggtggc taactacggc tacactagaa ggacagtatt tggatatctg gctctgctga 6600
agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaaaa accaccgctg 6660
gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaa ggatctcaag 6720
aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaac tcacgttaag 6780
ggatttttgt catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat 6840
gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct 6900
taatcagtga ggcacctatc tcagcgatct gtctatctcg ttcacctata gttgcctgac 6960
tccccgctcg gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa 7020
tgataccgcg agaccacgc tcaccggctc cagatttacc agcaataaac cagccagccg 7080
gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt 7140
gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca 7200
ttgctgcagg catcgtggtg tcacgctcgt cgtttgggtat ggcttcattc agctccgggt 7260
cccaacgatc aaggcgagtt acatgatccc ccatgttggt caaaaaagcg gttagctcct 7320
tcggctctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg 7380
cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactgggtg 7440
agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg 7500
cgtcaacacg ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa 7560
aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt 7620
aaccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt 7680
gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aaggcgaca cggaaatgtt 7740
gaatactcat actcttctt tttcaatatt attgaagcat ttatcagggt tattgtctca 7800
tgagcgggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat 7860
ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata 7920
aaaataggcg tatcacgagg ccctttcgtc ttcaagaatt gtt 7963

```

&lt;210&gt; 38

&lt;211&gt; 7966

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 38

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgcgt ttttgggtgt cgccggtgta cacaggaagt 240
gacaattttc gcgcgggttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360

```

actcatagcg	cgtaatat	gtctagggcc	gcggggactt	tgaccgttta	cgtggagact	420
cgcccaggtg	tttttctcag	gtgttttccg	cgttccgggt	caaagttggc	gttttattat	480
tatagtcagt	acgtaccagt	gcactggcct	agagcggccc	cattgcatac	gttgtatcca	540
tatcataata	tgtacattta	tattggctca	tgtccaacat	taccgccatg	ttgacattga	600
ttattgacta	gttattaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	660
gagttccgcg	ttacataact	tacggtaaat	ggcccgcctg	gctgaccgcc	caacgacccc	720
cgcccattga	cgtcaataat	gacgtatgtt	cccatagtaa	cgccaatagg	gactttccat	780
tgacgtcaat	gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	840
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggcccg	ctggcattat	900
gcccagtaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcac	960
gctattacca	tgggtgatgcg	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	1020
tcacggggat	ttccaagtct	ccaccccat	gacgtcaatg	ggagtttgtt	ttggcaccaa	1080
aatcaacggg	actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggcggt	1140
aggcgtgtac	gggtgggaggt	ctatataagc	agagctcggt	tagtgaaccg	tcagatcgcc	1200
tggagacgcc	atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	1260
cgtcaccgtc	gtcgacacgt	gtgatcagat	atcgcgggccg	ctctagaatg	cgcgatgaagg	1320
agaagtacca	gcacctgtgg	cgctggggct	ggcgtgggg	caccatgctg	ctgggcatgc	1380
tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgacacaagt	gtgggcccacc	cacgcctgcg	tgcccaccga	cccccaaccc	caggagggtg	1560
tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatgggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgctggaag	ctgaccccc	1680
tgtgctgag	cctgaagtgc	accgacctga	agaacgacac	caacaccaac	agcagcagcg	1740
gccgcatgat	catggagaag	ggcgagatca	agaactgcag	cttcaacatc	agcaccagca	1800
tccgcggcaa	ggtgcagaag	gagtacgcct	tcttctacaa	gctggacatc	atccccatcg	1860
acaacgacac	caccagctac	agcctgacca	gctgcaacac	cagcgtgatc	accaggcct	1920
gcccgaaggt	gagcttcgag	cccatcccca	accactactg	cgcccccgcc	ggcttcgcca	1980
tcctgaagtg	caaggacaag	aagttcaacg	gcaaggggccc	ctgcaccaac	gtgagcaccg	2040
tgacgtgcac	ccacggcatc	cgccccgtgg	tgagcaccga	gctgctgctg	aacggtagcc	2100
tggccgagga	ggaggtgggtg	atccgcagcg	ctaacttcgc	cgacaacgcc	aaggtgatca	2160
tcgtgcagct	gaacgagagc	gtggagatca	actgcacccg	ccccacaac	aacacccgca	2220
agagcatcca	catcgggccc	ggcgcgcct	tctacaccac	cggcgagatc	atcggcgaca	2280
tccgccaggc	ccactgcaac	ctgagccgcg	ccaagtggaa	cgacaccctg	aacaagatcg	2340
tgatcaagct	gcgcgagcag	ttcggaaca	agaccatcgt	gttcaagcac	agcagcggcg	2400
gcgaccccg	gatcgtgacc	cacagcttca	actgcggcgg	cgagttcttc	tactgcaaca	2460
gcgctagcct	gccctgccgc	atcaagcaga	tcatcaacat	gtggcagaag	gtgggcaagg	2520
ccatgtacgc	cccccccatc	agcggccaga	tccgctgcag	cagcaacatc	accggcctgc	2580
tgctgacccg	cgacggcggc	aacagcaaca	acgagagcga	gatcttccgt	ccgggcggcg	2640
gcgacatgcg	cgacaactgg	cgcagcgagc	tgtacaagta	caaggtgggtg	aagatcgagc	2700
ccctgggctg	ggccccacc	aaggccaagc	ttaccgtcca	ggcccgcag	ctgctgagcg	2760
gcacgtgca	gcagcagaac	aacctgctgc	gcgccatcga	ggcccagcag	cacctgctgc	2820
agctgaccgt	gtggggcatc	aagcagctgc	aggcccgcac	cctggccgtg	gagcgtacc	2880
tgaaggacca	gcagcagctc	gagcagatct	ggaaccacac	cacctggatg	gagtgggacc	2940
gcgagatcaa	caactacacc	agcctgatcc	acagcctgat	cgaggagagc	cagaaccagc	3000
aggagaagaa	cgagcaggag	ctgctggagc	tggacaagtg	ggccagcctg	tggaaactgg	3060
tcaacatcac	caactggctg	tgggtgaggat	ccagatctgc	tgtgccttct	agttgccagc	3120
catctgttgt	ttgcccctcc	cccgtgcctt	ccttgaccct	ggaagggtgc	actcccactg	3180
tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctatct	3240
tgggggggtgg	gggtggggcag	cacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	3300
ctggggatgc	gggtgggctct	atgggtaccc	agggcgcgat	aacttcgtat	aatgtatgct	3360
atacgaagtt	ataagatctg	tactgaaatg	tgtgggcgtg	gcttaagggt	gggaaagaat	3420
atataagggtg	gggtgcttat	gtagttttgt	atctgttttg	cagcagccgc	cgccgccatg	3480
agcaccact	cgtttgatgg	aagcattgtg	agctcatatt	tgacaacgcg	catgccccca	3540
tgggcccggg	tgcgtcagaa	tgtgatgggc	tccagcattg	atggtcgccc	cgtcctgccc	3600
gcaaactcta	ctaccttgac	ctacgagacc	gtgtctggaa	cgccgttgga	gactgcagcc	3660

tccgccgccc cttcagccgc tgcagccacc gcccgcgcca ttgtgactga ctttgctttc 3720  
ctgagccccc ttgcaagcag tgcagcttcc cgttcacccc cccgcgatga caagttgacg 3780  
gctcttttgg cacaatttga ttcttttgacc cgggaactta atgtcgtttc tcagcagctg 3840  
ttggatctgc gccagcaggt ttctgccttg aaggcttccc cccctcccaa tgcggtttaa 3900  
aacataaata aaaaaccaga ctctgttttg atttggatca agcaagtgtc ttgctgtctt 3960  
tatttagggg ttttgcgcgc gcggtaggcc cgggaccagc ggtctcggtc gttgagggtc 4020  
ctgtgtattt tttccaggac gtggtaaagg tgactctgga tgttcagata catgggcata 4080  
agcccgcttc tggggtggag gtagcaccac tgcagagctt catgctgcgg ggtggtggtg 4140  
tagatgatcc agtcgtagca ggagcgtcgg gcgtggtgcc taaaaatgtc tttcagtagc 4200  
aagctgattg ccaggggcag gcccttggtg taagtgttta caaagcgggt aagctgggat 4260  
gggtgcatac gtggggatat gagatgcata ttggactgta tttttagggt ggctatgttc 4320  
ccagccatat cctccggggg attcatgttg tgcagaacca ccagcacagt gtatccgggtg 4380  
cacttgggaa atttgtcatg tagcttagaa ggaaatgcgt ggaagaactt ggagacgccc 4440  
ttgtgacctc caagattttc catgcattcg tccataatga tggcaatggg cccacggggc 4500  
gcggcctggg cgaagatatt tctgggatca ctaacgtcat agttgtgttc caggatgaga 4560  
tcgtcatagg ccattttttac aaagcgcggg cggagggtgc cagactgcgg tataatggtt 4620  
ccatccggcc caggggcgta gttaccctca cagatttgca tttcccacgc tttgagttca 4680  
gatgggggga tcatgtctac ctgcggggcg atgaagaaaa cgttttccgg ggtaggggag 4740  
atcagctggg aagaaagcag gttcctgagc agctgcgact taccgcagcc ggtgggcccg 4800  
taaatcacac ctattaccgg ctgcaactgg tagttaagag agctgcagct gccgtcatcc 4860  
ctgagcaggg gggccacttc gttaagcatg tccctgactc gcattgtttc cctgaccaa 4920  
tccgccagaa ggcgctcgcc gccagcgat agcagttctt gcaaggaagc aaagtttttc 4980  
aacggtttga gaccgtccgc cgtaggcatg cttttgagcg tttgaccaag cagttccagg 5040  
cgggtcccaca gtcggtcac ctgctctacg gcactctgat ccagcatatc tctcgtttc 5100  
gcgggttggg gcggctttcg ctgtacggca gtagtcgggt ctcgtccaga cgggccaggg 5160  
tcatgtcttt ccacgggcgc agggctctcg tcagcgtagt ctgggtcacg gtgaaggggt 5220  
gcgtccggg ctgcgcgctg gccagggtgc gcttgaggct ggtcctgctg gtgctgaagc 5280  
gctgccggtc ttgcctctgc gcgtcggcca ggtagcattt gaccatggtg tcatagtcca 5340  
gccccccgc ggcgtggccc ttggcgcgca gcttgccctt ggaggaggcg ccgcacgagg 5400  
ggcagtgag acttttgagg gcgtagagct tgggcgcgag aaataccgat tccggggagt 5460  
aggcatccgc gccgcaggcc ccgcagacgg tctcgcattc cagagccag gtgagctctg 5520  
gccgttcggg gtcaaaaacc aggtttcccc catgcttttt gatgcgtttc ttacctctg 5580  
tttccatgag ccggtgtcca cgctcgggtg cgaagaggct gtccgtgtcc ccgtatacag 5640  
acttgagagg cctgtcctcg agcgggtgtc cgcggctctc ctgctataga aactcggacc 5700  
actctgagac aaaggctcgc gtccaggcca gcacgaagga ggctaagtgg gaggggtagc 5760  
ggtcgttgtc cactaggggg tccactcgct ccagggtgtg aagacacatg tcgccctctt 5820  
cggcatcaag gaaggtgatt ggtttgtagg tgtaggccac gtgaccgggt gttcctgaag 5880  
gggggctata aaagggggtg ggggcgcgtt cgtcctcact ctcttccgca tcgctgtctg 5940  
cgagggccag ctgttgggggt gagtcgacgc gaggtggat ggccttcccc attatgattc 6000  
ttctcgttcc cggcggcatc gggatgcccg cgttgaggcc catgctgtcc aggcaggtag 6060  
atgacgacca tcagggacag cttcaaggcc agcaaaaggc caggaaccgt aaaaaggccg 6120  
cgttgctggc gtttttccat aggtcccgcc cccctgacga gcacacaaa aatcgacgct 6180  
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 6240  
gtccctcgt gcgtctctct gttccgacct tgcgcttac cggatacctg tccgcctttc 6300  
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 6360  
aggctcgttc ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 6420  
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg 6480  
cagcagccac tggtaaacagg attagcagag cgagggtatgt aggcgggtgct acagagttct 6540  
tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 6600  
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 6660  
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 6720  
aagaagatcc tttgatcttt tctacgggggt ctgacgctca gtggaacgaa aactcacgtt 6780  
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaatataa 6840  
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 6900  
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagttgcct 6960



```

gactccccgt cgtgtagata actacgatac gggagggcctt accatctggc cccagtgtctg 7020
caatgatacc gcgagaccca cgctcaccgg ctccagatctt atcagcaata aaccagccag 7080
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 7140
attggtgccc ggaagctaga gtaagtagtt cgccagttaa tagtttgccc aacgttggtg 7200
ccattgtctg aggcacgtg gtgtcacgct cgtcgttttg tatggcttca ttcagctccg 7260
gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct 7320
ccttcggtec tccgacgtt gtcagaagta agttggccgc agtggttatca ctcatgggta 7380
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 7440
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 7500
cggcgtcaac acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 7560
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 7620
tgtaaccac tctgtcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 7680
ggtgagcaaa aacaggaagg caaatgccg caaaaaaggg aataaggcg acacggaaat 7740
gttgaatact catactcttc ctttttcaat attattgaag catttatcag gggtattgtc 7800
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 7860
catttccccg aaaagtgcc cctgacgtct aagaaaccat tattatcatg acattaacct 7920
ataaaaatag gcgtatcacg aggcctttc gtcttcaaga attgtt 7966

```

&lt;210&gt; 39

&lt;211&gt; 7870

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 39

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatct gtctaggggc gcggggactt tgaccgttta cgtggagact 420
cgcccagggt tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgcaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gccaggtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcac 960
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgcc cattgacgca aatgggcggg 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcacgctc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgccacc gagaagctgt gggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtgggcccac cagcctgcg tgcccaccga cccaacccc caggagggtg 1560

```

tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacgcta	gcaccagctg	caacaccagc	gtgatcacc	1740
aggcctgccc	caaggtgagc	ttcgagccca	tccccaacca	ctactgcgcc	cccgcggct	1800
tcgccatcct	gaagtgcaag	gacaagaagt	tcaacggcaa	gggcccctgc	accaacgtga	1860
gcaccgtgca	gtgcacccac	ggcatccgcc	ccgtggtgag	caccagctg	ctgctgaacg	1920
gtagcctggc	cgaggaggag	gtggtgatcc	gcagcgctaa	cttcgccgac	aacgccaagg	1980
tgatcatcgt	gcagctgaac	gagagcgtgg	agatcaactg	caccgcgcc	aacaacaaca	2040
cccgaagag	catccacatc	ggccccggcc	gcgccttcta	caccaccggc	gagatcatcg	2100
gcgacatccg	ccaggcccac	tgcaacctga	gccgcgccaa	gtggaacgac	accctgaaca	2160
agatcgtgat	caagctgcgc	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2220
gcggcggcga	ccccgagatc	gtgacccaca	gcttcaactg	cggcggcgag	ttctttctact	2280
gcaacagcac	ccagctgttc	aacagcacct	ggttcaacag	cacctggagc	accgagggca	2340
gcaacaacac	cgagggcagc	gacaccatca	ccctgccttg	ccgcatacag	cagatcatca	2400
acatgtggca	gaaggtgggc	aaggccatgt	acgccccccc	catcagcggc	cagatccgct	2460
gcagcagcaa	catcaccggc	ctgctgctga	cccgcgacgg	cggcaacagc	aacaacgaga	2520
gcgagatctt	ccgtccgggc	ggcggcgaca	tgcgcgacaa	ctggcgagc	gagctgtaca	2580
agtacaaggt	ggtgaagatc	gagccccctg	gcgtggcccc	caccaaggcc	aagcttaccg	2640
tccaggcccc	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	ctgcgcgcca	2700
tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	ctgcaggccc	2760
gcaccctggc	cgtggagcgc	tacctgaagg	accagcagca	gctcgagcag	atctggaacc	2820
acaccacctg	gatggagtgg	gaccgcgaga	tcaacaacta	caccagcctg	atccacagcc	2880
tgatcgagga	gagccagaac	cagcaggaga	agaacgagca	ggagctgctg	gagctggaca	2940
agtgggcccag	cctgtggaac	tgggtcaaca	tcaccaactg	gctgtggtga	ggatccagat	3000
ctgctgtgcc	ttctagtgtc	cagccatctg	ttgtttgccc	ctccccctg	ccttccttga	3060
ccctggaagg	tgccactccc	actgtccttt	cctaataaaa	tgaggaaatt	gcatcgcat	3120
gtctgagtag	gtgtcattct	attctggggg	gtgggggtgg	gcagcacagc	aagggggagg	3180
attgggaaga	caatagcagg	catgctgggg	atgcgggtgg	ctctatgggt	acccagggcc	3240
gcataacttc	gtataatgta	tgctatacga	agttataaga	tctgtactga	aatgtgtggg	3300
cgtggcctta	gggtgggaaa	gaatatataa	gggtgggggtc	ttatgtagtt	ttgtatctgt	3360
tttgccagcag	ccgccgcgcg	catgagcacc	aactcgtttg	atggaagcat	tgtgagctca	3420
tatttgacaa	cgcgcagtcg	cccattgggc	gggggtgcgtc	agaatgtgat	gggctccagc	3480
attgatggtc	gccccgtcct	gcccgcacaa	tctactacct	tgacctacga	gaccgtgtct	3540
ggaacgccgt	tggagactgc	agcctccgcc	gccgcttcag	ccgctgcagc	caccgcccgc	3600
gggatttgtga	ctgactttgc	tttcctgagc	ccgcttgcaa	gcagtgcagc	ttcccgttca	3660
tccgcccgcg	atgacaagtt	gacggctcct	ttggcacaa	tggattcttt	gaccggggaa	3720
cttaatgtcg	ttctctcagca	gctgttggat	ctgcgccagc	aggtttctgc	cctgaaggct	3780
tcctcccttc	ccaatgcggg	ttaaaacata	aataaaaaaac	cagactctgt	ttggatttgg	3840
atcaagcaag	tgtcttgctg	tctttattta	gggtttttgc	gcgcgcggta	ggcccgggac	3900
cagcggcttc	ggtcgttgag	ggtcctgtgt	attttttcca	ggacgtggta	aaggtgactc	3960
tggatgttca	gatacatggg	cataagcccc	tctctggggg	ggaggtagca	ccactgcaga	4020
gcttcagctg	gcgggggtgg	gttgtagatg	atccagtcgt	agcaggagcg	ctgggcgtgg	4080
tgcctaaaaa	tgtctttcag	tagcaagctg	attgccaggg	gcaggccctt	ggtgtaagtg	4140
tttacaagc	ggttaagctg	ggatgggtgc	atacgtgggg	atatgagatg	catcttggac	4200
tgtattttta	ggttggttat	gttcccagcc	atatccctcc	ggggattcat	gttgtgcaga	4260
accaccagca	cagtgtatcc	ggtgcacttg	ggaaatttgt	catgtagctt	agaaggaaat	4320
gcgtggaaga	acttgagagc	gcccttgtga	cctccaagat	tttccatgca	ttcgtccata	4380
atgatggcaa	tgggcccacg	ggcggcgggc	tgggcgaaga	tatttctggg	atcactaacg	4440
tcatagtgtg	gttccaggat	gagatcgtca	taggccattt	ttacaaagcg	cgggcggagg	4500
gtgccagact	gcggtataat	ggttccatcc	ggcccagggg	cgtagtacc	ctcacagatt	4560
tgcattttcc	acgctttgag	ttcagatggg	gggatcatgt	ctacctgcgg	ggcgatgaag	4620
aaaacggttt	ccggggtagg	ggagatcagc	tgggaagaaa	gcaggttcct	gagcagctgc	4680
gacttaccgc	agccggtggg	cccgtaaatc	acacctatta	ccggctgcaa	ctggtagtta	4740
agagagctgc	agctgcgctc	atccctgagc	agggggggcca	cttcgttaag	catgtccctg	4800
actcgcatgt	tttccctgac	caaatccgcc	agaaggcgct	cgcgcgccag	cgatagcagt	4860

tcttgcaagg aagcaaagtt tttcaacggt ttgagaccgt ccgccgtagg catgcttttg 4920  
agcgtttgac caagcagttc caggcgggtcc cacagctcgg tcacctgctc tacggcatct 4980  
cgatccagca tatctcctcg tttcgcgggt tggggcggct ttcgctgtac ggcagtagtc 5040  
ggctgctcgtc cagacgggccc aggggtcatgt ctttccacgg gcgcagggtc ctgctcagcg 5100  
tagtctgggt cacgggtgaag ggggtgcgtc cgggctgcgc gctggccagg gtgcgcttga 5160  
ggctggctct gctgggtgctg aagcgctgcc ggtcttcgcc ctgcgcgtcg gccaggtagc 5220  
atltgaccat ggtgtcatag tccagccccct ccgcggcgtg gcccttggcg cgcagcttgc 5280  
ccttgaggga ggcgccgcac gaggggcagt gcagactttt gagggcgtag agcttgggcg 5340  
cgagaaatac cgattccggg gagtaggcat ccgcgccgca ggccccgcag acggtctcgc 5400  
attccacgag ccagggtgagc tctggccggt cgggggtcaaa aaccagggtt ccccatgct 5460  
ttttgatgcg tttcttacct ctggtttcca tgagccgggtg tccacgctcg gtgacgaaaa 5520  
ggctgtccgt gtccccgtat acagacttga gaggcctgtc ctgcagcgggt gttccgcgggt 5580  
cctcctcgtc tagaaactcg gaccactctg agacaaaggc tcgcgtccag gccagcacga 5640  
aggaggctaa gtgggagggg tagcgggtcgt tgtccactag ggggtccact cgctccaggg 5700  
tgtgaagaca catgtcgcgc tcttcggcat caaggaaggt gattggtttg taggtgtagg 5760  
ccacgtgacc ggggtgttct gaagggggggc tataaaaggg ggtggggggcg cgttcgtcct 5820  
cactctcttc cgcctcgtcg tctgcgaggg ccagctgttg ggggtgagtc acgcgagggt 5880  
ggatggcctt ccccatatg attctctctg ctccggcgcg catcgggatg cccgcgttgc 5940  
aggccatgct gtccaggcag gtagatgacg accatcaggg acagcttcaa ggccagcaaa 6000  
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 6060  
acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaaccgcaca ggactataaa 6120  
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tctgttccg accctgccgc 6180  
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 6240  
gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 6300  
ccccggttca gcccgaccgc tgcgccttat ccggttaacta tcgtcttgag tccaaccgg 6360  
taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgagg 6420  
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 6480  
cagtatttgg tatctgcgtc ctgctgaagc cagttacctt cggaaaaaga gttggtagct 6540  
cttgatccgg caaacaacc accgctggta gcggtggtt ttttgtttgc aagcagcaga 6600  
ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctcgacg 6660  
ctcagtgga cgaaaaactca cgttaaggga ttttgggtcat gagattatca aaaaggatct 6720  
tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 6780  
aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 6840  
tatttctgtc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg 6900  
gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag 6960  
atltatcagc aataaaccag ccagccggaa gggccgagcg cagaagtgggt cctgcaactt 7020  
tatccgcctc catccagctc attaatgtt gccgggaagc tagagtaagt agttcgccag 7080  
ttaatagttt gcgcaacggt gttgccattg ctgcaggcat cgtgggtgtca cgctcgtcgt 7140  
ttgggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 7200  
tgttgtgcaa aaaagcgggt agctccttcg gtccctccgat cgttgtcaga agtaagtgg 7260  
ccgcagtggt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 7320  
ccgtaagatg cttttctgtg actgggtgagt actcaaccaa gtcattctga gaatagtgt 7380  
tgccggcacc gagttgctct tgcggggcgt caacacggga taataccgcg ccacatagca 7440  
gaactttaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 7500  
taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 7560  
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaat gccgcaaaaa 7620  
agggaataag ggcgacacgg aaatggtgaa tactcatact cttccttttt caatattatt 7680  
gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 7740  
ataaacaat aggggttccg cgcacatttc ccgaaaagt gccacctgac gtctaagaaa 7800  
ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc 7860  
aagaattggt 7870



&lt;211&gt; 7906

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 40

```

ttaaattaacc gcaatttctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggtgtg tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccagggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgacctatg ggactttcct acttggcagt acatctacgt attagtcac 960
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgaggata gcggtttgac 1020
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtagagg 1320
agaagtacca gcacctgtgg cgtgggggt ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgcacc gagagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtgggcccacc cagcctgcg tgcaccacca cccaacccc caggagggtg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtagag ctgaccccc 1680
tgtgcgtag cctgaagtgc accgacgcta gcaagaactg cagcttcaac atcagacca 1740
gcatccgcg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 1800
tcgacaacga caccaccagc tacagcctga ccagctgcaa caccagcgtg atcaccagg 1860
cctgcccga ggtgagcttc gagcccatcc ccaaccacta ctgcgcccc gccggcttcg 1920
ccatcctgaa gtgcaaggac agaagttca acggcaaggg cccctgcacc aacgtgagca 1980
ccgtgcagtg caccacggc atccgcccc tgggtgagcac ccagctgctg ctgaacggta 2040
gcctggccga ggaggagggt gtgatccgca gcgctaactt cgccgacaac gccaagggtg 2100
tcatcgtgca gctgaacgag agcgtggaga tcaactgcac ccgcgctagc gccactgca 2160
acctgagccg cgccaagtgg aacgacaccc tgaacaagat cgtgatcaag ctgcgcgagc 2220
agttcggcaa caagaccatc gtgttcaagc acagcagcgg cggcgacccc gagatcgtga 2280
cccacagctt caactgcggc ggcgagttct tctactgcaa cagcaccag ctgttcaaca 2340
gcacctggtt caacagcacc tggagcaccg agggcagcaa caacaccgag ggcagcgaca 2400
ccatcacctt gccctgccgc atcaagcaga tcatcaacat gtggcagaag gtgggcaagg 2460
ccatgtacgc ccccccatc agcggccaga tccgctgcag cagcaacatc accggcctgc 2520
tgctgacccg cgacggcggc aacagcaaca acgagagcga gatcttccgt ccgggcggcg 2580
gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg aagatcgagc 2640
ccctgggcgt ggccccacc aaggccaagc ttaccgtcca ggcccgcag ctgctgagcg 2700
gcatcgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag cacctgctgc 2760
agctgaccgt gtggggcacc aagcagctgc agggccgcac cctggccgtg gagcgctacc 2820

```

tgaaggacca	gcagcagctc	gagcagatct	ggaaccacac	cacctggatg	gagtgggacc	2880
gcgagatcaa	caactacacc	agcctgatcc	acagcctgat	cgaggagagc	cagaaccagc	2940
aggagaagaa	cgagcaggag	ctgctggagc	tggacaagtg	ggccagcctg	tggaaactggt	3000
tcaacatcac	caactggctg	tggtaggat	ccagatctgc	tgtgccttct	agttgccagc	3060
catctgttgt	ttgcccctcc	cccgctgcctt	ccttgacctt	ggaagggtgcc	actcccactg	3120
tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	3180
tgggggggtg	ggtggggcag	cacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	3240
ctggggatgc	ggtgggctct	atgggtaccc	agggccgcac	aacttcgtat	aatgtatgct	3300
atacgaagtt	ataagatctg	tactgaaatg	tgtgggcgtg	gcttaagggg	gggaaagaat	3360
atataaggtg	ggggtcttat	gtagttttgt	atctgttttg	cagcagccgc	cgccgccatg	3420
agcaccaact	cgtttgatgg	aagcattgtg	agctcatatt	tgacaacgcg	catgccccca	3480
tggggccggg	tgcgtcagaa	tgtgatgggc	tccagcattg	atggtcgccc	cgtcctgccc	3540
gcaaactcta	ctaccttgac	ctacgagacc	gtgtctggaa	cgccgttggg	gactgcagcc	3600
tccgcccggc	cttcagccgc	tgcagccacc	gcccgcggga	ttgtgactga	ctttgctttc	3660
ctgagcccg	ttgcaagcag	tgcagcttcc	cgttcatccg	cccgcgatga	caagttgacg	3720
gctcttttgg	cacaattgga	ttcttttgacc	cgggaactta	atgtcgtttc	tcagcagctg	3780
ttggatctgc	gccagcaggt	ttctgcccctg	aaggcttccct	cccctcccaa	tgcggtttaa	3840
aacataaata	aaaaaccaga	ctctgttttg	atltggatca	agcaagtgtc	ttgctgtctt	3900
tatttagggg	ttttgcgcgc	gcggtagggc	cgggaccagc	ggtctcgggtc	gttgaggggtc	3960
ctgtgtattt	tttccaggac	gtggtaaagg	tgactctgga	tgttcagata	catgggcata	4020
agcccgtctc	tgggggtggag	gtagcaccac	tgcagagctt	catgctgcgg	ggtggtgttg	4080
tagatgatcc	agtcgtagca	ggagcgcctg	gcgtgggtgc	taaaaatgtc	tttcagtagc	4140
aagctgattg	ccaggggcag	gcccttggtg	taagtgttta	caaagcgggt	aagctgggat	4200
gggtgcatac	gtggggatat	gagatgcac	ttggactgta	tttttaggtt	ggctatgttc	4260
ccagccatat	ccctccgggg	attcatgttg	tgcagaacca	ccagcacagt	gtatccgggtg	4320
cacttgggaa	atttgtcatg	tagcttagaa	ggaaatgcgt	ggaagaactt	ggagacgccc	4380
ttgtgacctc	caagattttc	catgcattcg	tccataatga	tggcaatggg	cccacggggc	4440
gcggcctggg	cgaagatatt	tctgggatca	ctaactgcac	agttgtgttc	caggatgaga	4500
tcgtcatagg	ccattttttac	aaagcgcggg	cggaggggtc	cagactgcgg	tataatgggt	4560
ccatccggcc	cagggggcgta	gttaccctca	cagatttgca	tttcccacgc	tttgagttca	4620
gatgggggga	tcatgtctac	ctgcggggcg	atgaagaaaa	cggtttccgg	ggtaggggag	4680
atcagctggg	aagaaagcag	gttcctgagc	agctgcgact	taccgcagcc	ggtgggccc	4740
taaatcacac	ctattaccgg	ctgcaactgg	tagttaagag	agctgcagct	gccgtcatcc	4800
ctgagcaggg	gggcccacttc	gttaagcatg	tccctgactc	gcatgttttc	cctgacccaa	4860
tccgccagaa	ggcgcctgcg	gcccagcgat	agcagttctt	gcaaggaagc	aaagtttttc	4920
aacggtttga	gaccgtccgc	cgtaggcatg	cttttgagcg	tttgaccaag	cagttccagg	4980
cgggtcccaca	gctcgggtcac	ctgctctacg	gcatctcgat	ccagcatatc	tcctcgtttc	5040
gcgggttggg	gcggcttttcg	ctgtacggca	gtagtccgtg	ctcgtccaga	cggggccagg	5100
tcatgtcttt	ccacggggcg	agggtccctg	tcagcgtagt	ctgggtcacg	gtgaaggggt	5160
gcgctccggg	ctgcgcgctg	gccaggggtg	gcttgaggct	ggtcctgctg	gtgctgaagc	5220
gctgccgggt	ttcgccctgc	gcgtcggcca	ggttagcattt	gaccatgggt	tcatagtcca	5280
gccccctccg	ggcgtggccc	ttggcgcgca	gcttgccctt	ggaggaggcg	ccgcacgagg	5340
ggcagtgacg	acttttgagg	gcgtagagct	tgggcgcgag	aaataccgat	tccggggagt	5400
aggcatccgc	gccgcaggcc	ccgcagacgg	tctcgcattc	cacgagccag	gtgagctctg	5460
gccgttcggg	gtcaaaaacc	aggtttcccc	catgcttttt	gatgcgtttc	ttacctctgg	5520
tttccatgag	ccggtgtcca	cgctcgggtg	cgaaaaggct	gtccgtgtcc	ccgtatacag	5580
acttgagagg	cctgtcctcg	agcgggtgtc	cgcggtcctc	ctcgtataga	aactcggacc	5640
actctgagac	aaaggctcgc	gtccaggcca	gcacgaagga	ggctaagtgg	gaggggtagc	5700
ggtcgttgtc	cactaggggg	tccactcgct	ccagggtgtg	aagacacatg	tcgcccctct	5760
cggcatcaag	gaagggtgatt	ggttttgtagg	tgtaggccac	gtgaccgggt	gttcctgaag	5820
gggggctata	aaaggggggtg	ggggcgcggt	cgtcctcact	ctcttccgca	tcgctgtctg	5880
cgagggccag	ctgttgggggt	gagtcgacgc	gaggctggat	ggccttcccc	attatgatcc	5940
ttctcgcctc	cggcggcatc	gggatgcccc	cgttgacggc	catgctgtcc	aggcaggtag	6000
atgacgacca	tcagggacag	cttcaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	6060
cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatacaaaa	aatcgacgct	6120

```

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 6180
gctccctcgt gcgctctcct gtcccgaccc tgccgcttac cggataacctg tccgcctttc 6240
tcccttcggg aagcgtggcg cttctctcata gctcacgctg taggtatctc agttcgggtg 6300
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 6360
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tgcgactgg 6420
cagcagccac tggtaacagg attagcagag cgaggatgtt aggcgggtgt acagagttct 6480
tgaagtgggtg gcctaactac ggctacacta gaaggacagt atttgggtatc tgcgctctgc 6540
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 6600
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 6660
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 6720
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 6780
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 6840
gcttaatcag tgaggcacct atctcagcga tctgtctatt tctgtcatcc atagttgcct 6900
gactccccgt cgtgtagata actacgatac gggagggtt accatctggc cccagtgcgt 6960
caatgatacc gcgagacca cgctcaccgg ctccagattt atcagcaata aaccagccag 7020
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 7080
attggtgccc ggaagctaga gtaagtagtt cgccagttaa tagtttgccg aacgttggtg 7140
ccattgctgc aggcacgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 7200
gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct 7260
ccttcgggtc tccgatcgtt gtcagaagta agttggccgc agtggtatca ctcattggtt 7320
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 7380
gtgagtactc aaccaagtca ttctgagaat agtgatgctg gcgaccgagt tgctcttgcc 7440
cggcgtcaac acgggataat accgcgccac atagcagaac tttaaaagtg ctcattcattg 7500
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 7560
tgtaaccac tctgtcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 7620
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 7680
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 7740
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca 7800
catttccccg aaaagtgcc cctgacgtct aagaaaccat tattatcatg acattaacct 7860
ataaaaatag gcgtatcacg aggccttttc gtcttcaaga attgtt 7906

```

&lt;210&gt; 41

&lt;211&gt; 7909

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 41

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgt cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatct gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagtggc gttttattat 480
tatagtcagt acgtaccagt gcaactggct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780

```



tgacgtcaat	gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	840
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggcccgc	ctggcattat	900
gcccagtaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcate	960
gctattacca	tggatgatgcg	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	1020
tcacggggat	ttccaagtct	ccacccccatt	gacgtcaatg	ggagtgtgtt	ttggcaccaa	1080
aatcaacggg	actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggagggt	1140
aggcgtgtac	ggtgggagggt	ctatataagc	agagctcggt	tagtgaaccg	tcagatcgcc	1200
tggagacgcc	atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	1260
cgtcaccgtc	gtcgacacgt	gtgatcagat	atcgcgcccg	ctctagaatg	cgcgtagaagg	1320
agaagtacca	gcacctgtgg	cgctggggct	ggcgtggggg	caccatgctg	ctgggcatgc	1380
tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgcacaacgt	gtggggccacc	cacgcctgcg	tgcccaccga	ccccaaacccc	caggagggtgg	1560
tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatgggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacgcta	gcaagaactg	cagcttcaac	atcagcacca	1740
gcatccgcgg	caagggtgcag	aaggagtagc	ccttcttcta	caagctggac	atcatcccca	1800
tcgacaacga	caccaccagc	tacagcctga	ccagctgcaa	caccagcgtg	atcaccagg	1860
cctgccccaa	ggtgagcttc	gagcccatcc	ccaaccacta	ctgcgcccc	gccggcttcg	1920
ccatcctgaa	gtgcaaggac	aagaagttca	acggcaaggg	cccctgcacc	aacgtgagca	1980
ccgtgcagtg	cacccacggc	atccgccccg	tgggtgagcac	ccagctgctg	ctgaacggta	2040
gcctggccga	ggaggagggtg	gtgatccgca	gcgctaactt	cgccgacaac	gccaagggtga	2100
tcacgtgca	gctgaacgag	agcgtggaga	tcaactgcac	ccgccccaac	aacaacaccc	2160
gcaagagcat	ccacatcggc	ccgggcccgcg	ccttctacac	caccggcgag	atcatcggcg	2220
acatccgcca	ggcccactgc	aacctgagcc	gcgccaagtg	gaacgacacc	ctgaacaaga	2280
tcgtgatcaa	gctgcgcgag	cagttcggca	acaagaccat	cgtgttcaag	cacagcagcg	2340
gcggcgaccc	cgagatcggtg	accacagct	tcaactgcgg	cggcgagttc	ttctactgca	2400
acagcgtag	cctgccctgc	cgcatcaagc	agatcatcaa	catgtggcag	aagggtgggca	2460
aggccatgta	cgcccccccc	atcagcggcc	agatccgctg	cagcagcaac	atcaccggcc	2520
tgctgctgac	ccgcgacggc	ggcaacagca	acaacgagag	cgagatcttc	cgtccgggcg	2580
gcggcgacat	gcgcgacaac	tggcgcagcg	agctgtacaa	gtacaagggtg	gtgaagatcg	2640
agccctggg	cgtggcccc	accaaggcca	agcttaccgt	ccaggcccgc	cagctgctga	2700
gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggcccag	cagcacctgc	2760
tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccg	caccctggcc	gtggagcgct	2820
acctgaagga	ccagcagcag	ctcgagcaga	tctggaacca	caccacctgg	atggagtggg	2880
accgcgagat	caacaactac	accagcctga	tccacagcct	gatcgaggag	agccagaacc	2940
agcaggagaa	gaacgagcag	gagctgctgg	agctggacaa	gtgggcccagc	ctgtggaact	3000
ggttcaacat	caccaactgg	ctgtgggtgag	gatccagatc	tgtgtgtgct	tctagtgtgc	3060
agccatctgt	tgtttgcccc	tccccgtgc	cttccctgac	cctggaagggt	gccactccca	3120
ctgtcctttc	ctaataaaaat	gaggaaattg	catcgcatgt	tctgagtagg	tgtcattcta	3180
ttctgggggg	tgggggtggg	cagcacagca	agggggaggga	ttgggaagac	aatagcaggc	3240
atgctgggga	tgcgggtggc	tctatgggta	cccaggggccg	cataacttcg	tataatgtat	3300
gctatacgaa	gttataagat	ctgtactgaa	atgtgtgggc	gtggcttaag	gggtgggaaag	3360
aatatataag	gtgggggtct	tatgtagtgt	tgtatctgtt	ttgcagcagc	cgccgcccgc	3420
atgagcacca	actcgtttga	tggaaagcatt	gtgagctcat	atltgacaac	gcgcagtccc	3480
ccatggggccg	gggtgcgtca	gaatgtgatg	ggctccagca	ttgatgggtc	ccccgtcctg	3540
cccgcaaaact	ctactacctt	gacctacgag	accgtgtctg	gaacgccgtt	ggagactgca	3600
gcctccgccc	ccgcttcagc	cgctgcagcc	accgcccgcg	ggattgtgac	tgactttgct	3660
ttcctgagcc	cgcttgcaag	cagtgcagct	tcccgttcat	ccgcccgcga	tgacaagtgt	3720
acggctcttt	tggcacaatt	ggattctttg	acccgggaac	ttaatgtcgt	ttctcagcag	3780
ctgttggatc	tgcgccagca	ggtttctgcc	ctgaaggctt	cctccccctc	caatgcgggt	3840
taaaacataa	ataaaaaaac	agactctgtt	tggatttggga	tcaagcaagt	gtcttgctgt	3900
ctttatcttag	gggttttgcg	cgcgcggtag	gccggggacc	agcggctctg	gtcgttgagg	3960
gtcctgtgta	ttttttccag	gacgtggtaa	aggtgactct	ggatgttcag	atacatgggc	4020
ataagcccg	ctctgggggtg	gaggtagcac	cactgcagag	cttcatgctg	cggggtgggtg	4080

ttgtagatga	tccagtcgta	gcaggagcgc	tgggcgtggt	gcctaaaaat	gtctttcagt	4140
agcaagctga	ttgccagggg	caggcccttg	gtgtaagtgt	ttacaaagcg	gttaagctgg	4200
gatgggtgca	tacgtgggga	tatgagatgc	atcttggact	gtatttttag	gttggctatg	4260
ttcccagcca	tatccctccg	gggattcatg	ttgtgcagaa	ccaccagcac	agtgtatccg	4320
gtgcacttgg	gaaatttgtc	atgtagctta	gaaggaaatg	cgtggaagaa	cttggagacg	4380
cccttgtgac	ctccaagatt	ttccatgcat	tcgtccataa	tgatggcaat	gggcccacgg	4440
gcggcggcct	gggcgaagat	atttctggga	tcactaacgt	catagtgtgt	ttccaggatg	4500
agatcgtcac	aggccatttt	tacaaagcgc	gggcggaggg	tgccagactg	cgggtataatg	4560
gttccatccg	gcccaggggc	gtagttagcc	tcacagattt	gcatttccca	cgctttgagt	4620
tcagatgggg	ggatcatgtc	tacctgcggg	gcgatgaaga	aaacggtttc	cggggtaggg	4680
gagatcagct	gggaagaaag	caggttcctg	agcagctgcg	acttaccgca	gccggtgggc	4740
ccgtaaataca	cacctattac	cggctgcaac	tggtagttaa	gagagctgca	gctgccgtca	4800
ttcctgagca	ggggggccac	ttcgttaagc	atgtccctga	ctcgcatgtt	ttccttgacc	4860
aaatccgcca	gaaggcgctc	gccgcccgag	gatagcagtt	cttgcaagga	agcaaagttt	4920
ttcaacgggt	tgagaccgtc	cgcgtagggc	atgcttttga	gcgtttgacc	aagcagttcc	4980
aggcgggtcc	acagctcggg	cacctgctct	acggcatctc	gatccagcat	atctcctcgt	5040
ttcgcggggt	ggggcgggct	tcgctgtacg	gcagtagtcg	gtgctcgtcc	agacggggcca	5100
gggtcatgtc	tttccacggg	cgcagggtcc	tcgtcagcgt	agtctgggtc	acgggtgaagg	5160
gggtgcgtcc	gggctgcgcg	ctggccaggg	tgcgcttgag	gctgggtcctg	ctgggtgctga	5220
agcgtgccc	gtcttcgccc	tgcgcgtcgg	ccaggtagca	tttgaccatg	gtgtcatagt	5280
ccagcccctc	cgcggcgtgg	cccttgggcg	gcagcttgcc	cttgaggagg	gcgcgcgacg	5340
aggggcagtg	cagacttttg	agggcgtaga	gcttgggcgc	gagaaatacc	gattccgggg	5400
agtaggcac	cgcgcgcgag	gccccgcaga	cggctctcgca	ttccacgagc	cagggtgagct	5460
ctggccgttc	ggggtcaaaa	accagggttc	ccccatgctt	tttgatgcgt	ttcttacctc	5520
tggtttccat	gagccgggtg	ccacgctcgg	tgacgaaaag	gctgtccgtg	tccccgtata	5580
cagacttgag	aggcctgtcc	tcgagcgggt	ttccgcgggt	ctcctcgtat	agaaactcgg	5640
accactctga	gacaaaggct	cgcgtccagg	ccagcacgaa	ggaggctaag	tgggaggggt	5700
agcggtcgtt	gtccactagg	gggtccactc	gctccagggt	gtgaagacac	atgtcgccct	5760
cttcggcatc	aaggaagggt	attgggtttg	aggtgtaggc	cacgtgaccg	ggtgttcctg	5820
aagggggggt	ataaaagggt	gtggggggcg	gttcgtcctc	actctcttcc	gcacgcgtgt	5880
ctgcgagggc	cagctgttgg	ggtgagtcga	cgcgaggctg	gatggccttc	cccattatga	5940
ttcttctcgc	ttccggcggc	atcgggatgc	ccgcgttgca	ggccatgctg	tccaggcagg	6000
tagatgacga	ccatcaggga	cagcttcaag	gccagcaaaa	ggccagggaac	cgtaaaaagg	6060
ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	6120
gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	ataccaggcg	tttccccctg	6180
gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	6240
ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	ctgtagggtat	ctcagttcgg	6300
tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	6360
ggccttatc	cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	6420
tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggg	gctacagagt	6480
tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	6540
tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	6600
ccgctggtag	cgggtggttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	6660
ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	6720
gttaagggat	tttgggtcatg	agattatcaa	aaaggatcct	cacctagatc	cttttaaatt	6780
aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	6840
aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	6900
cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	6960
ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	7020
cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	7080
ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	7140
ttgccattgc	tgcaggcatc	gtggtgtcac	gctcgtcgtt	tggtatggct	tcattcagct	7200
ccggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcggtta	7260
gctccttcgg	tcctccgate	gttgtcagaa	gtaagtgggc	cgcagtggtta	tcactcatgg	7320
ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	7380

```

ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 7440
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7500
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 7560
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 7620
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 7680
aatggtgaat actcatactc ttccctttttc aatattattg aagcatttat cagggttatt 7740
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7800
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7860
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattggt 7909

```

&lt;210&gt; 42

&lt;211&gt; 7867

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 42

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatattt 60
ggattgaagc caatatgata atgaggggggt ggagtttbtg acgtggcgcg gggcggtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgt cgccgggtgt cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaa atbtggcgtaa ccgagtaaga 300
tttggccatt ttgcgggaa aactgaataa gaggaagtga aatctgaata attttbtgtt 360
actcatagcg cgtaatatatt gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccagggtg tttttctcag gtgttttccg cgttccgggt caaagtgtgc gttttattat 480
tatagtcagt acgtaccagt gcaactggcc agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcac 960
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcgggt 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tgagacgccc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctgggggt ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcggccacc gagaagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtggggccacc cagcctgctg tggccaccga cccaacccc caggaggtgg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctgtaag ctgaccccc 1680
tgtgcgtgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740
gccgcatgat catggagaag ggcgagatca agaactgcag cttcgctagc accagctgca 1800
acaccagcgt gatcaccag gcctgcccc aagtgagctt cgagcccatc cccaaccact 1860
actgcgcccc cgccggcttc gccatcctga agtgcaagga caagaagttc aacggcaagg 1920
gcccctgcac caacgtgagc accgtgcagt gcacccacgg catccgcccc gtggtgagca 1980
cccagctgct gctgaacgggt agcctggccg aggaggaggt ggtgatccgc agcgctaact 2040

```



tgcgcgacaa	cgccaagggtg	atcatcgtgc	agctgaacga	gagcgtggag	atcaactgca	2100
cccgcgctag	cgcccactgc	aacctgagcc	gcgccaagt	gaacgacacc	ctgaacaaga	2160
tcgtgatcaa	gctgcgcgag	cagttcggca	acaagaccat	cgtgttcaag	cacagcagcg	2220
gcggcgaccc	cgagatcgtg	acccacagct	tcaactgcgg	cggcgagttc	ttctactgca	2280
acagcaccca	gctgttcaac	agcacctgg	tcaacagcac	ctggagcacc	gagggcagca	2340
acaacaccga	gggcagcgac	accatcacc	tgccctgccg	catcaagcag	atcatcaaca	2400
tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	atccgctgca	2460
gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	aacgagagcg	2520
agatcttccg	tccgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	ctgtacaagt	2580
acaagggtgg	gaagatcgag	cccctgggcg	tggcccccac	caaggccaag	cttaccgtcc	2640
agggccgcca	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	cgcgccatcg	2700
agggccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	caggcccgcg	2760
ccctggccgt	ggagcgctac	ctgaaggacc	agcagcagct	cgagcagatc	tggaaccaca	2820
ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgac	cacagcctga	2880
tcgaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	ctggacaagt	2940
gggcccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtgagga	tccagatctg	3000
ctgtgccttc	tagttgccag	ccatctgttg	tttgccctc	ccccgtgcct	tccttgaccc	3060
tgaaggtgc	cactccact	gtcctttcct	aataaaatga	ggaaattgca	tcgcatgtgc	3120
tgagtaggtg	tcattctatt	ctgggggggtg	gggtggggca	gcacagcaag	ggggaggatt	3180
gggaagacaa	tagcaggcat	gctggggatg	cgggtgggctc	tatgggtacc	cagggccgca	3240
taacttcgta	taatgtatgc	tatacgaagt	tataagatct	gtactgaaat	gtgtgggcgt	3300
ggcttaaggg	tgggaaagaa	tatataaggt	gggggtctta	tgtagttttg	tatctgtttt	3360
gcagcagccg	ccgcgcgat	gagcaccac	tcgtttgatg	gaagcattgt	gagctcatat	3420
ttgacaacgc	gcatgcccc	atggggccggg	gtgcgtcaga	atgtgatggg	ctccagcatt	3480
gatggtcgcc	ccgtcctgcc	cgcaaactct	actaccttga	cctacgagac	cgtgtctgga	3540
acgccgttgg	agactgcagc	ctccgcgcgc	gcttcagccg	ctgcagccac	cgccgcggg	3600
attgtgactg	actttgcttt	cctgagcccg	cttgcaagca	gtgcagcttc	ccgttcaccc	3660
gcccgcgatg	acaagttgac	ggctcttttg	gcacaatttg	attctttgac	ccgggaactt	3720
aatgtcgttt	ctcagcagct	gttggatctg	cgccagcagg	tttctgccct	gaaggcttcc	3780
tcccctccca	atgcggttta	aaacataaat	aaaaaaccag	actctgtttg	gattttggatc	3840
aagcaagtgt	cttgctgtct	ttatlttaggg	gttttgcgcg	cgcggtaggc	ccgggaccag	3900
cggctctcgg	cgttgagggt	cctgtgtatt	ttttccagga	cgtggtaaag	gtgactctgg	3960
atgttcagat	acatgggcat	aagcccgctc	ctgggggtgga	ggtagacca	ctgcagagct	4020
tcatgctgcg	gggtgggtgt	gtagatgac	cagtcgtagc	aggagcgtg	ggcgtgggtg	4080
ctaaaaatgt	ctttcagtag	caagctgatt	gccaggggca	ggcccttggt	gtaagtgttt	4140
acaaagcggt	taagctggga	tgggtgcata	cgtggggata	tgagatgcat	cttggactgt	4200
attttttaggt	tggctatgtt	cccagccata	tccctccggg	gattcatgtt	gtgcagaacc	4260
accagcacag	tgtatccggt	gcacttgga	aatttgctcat	gtagcttaga	aggaaatgcg	4320
tggagaagact	tggagacgcc	cttgtgacct	ccaagatttt	ccatgcattc	gtccataatg	4380
atggcaatgg	gcccacgggc	ggcgccctgg	gcgaagatat	ttctgggatc	actaacgtca	4440
tagttgtgtt	ccaggatgag	atcgtcatag	gccattttta	caaagcgcg	gcggaggggtg	4500
ccagactgcg	gtataatggt	tccatccggc	ccaggggcgt	agttaccctc	acagatttgc	4560
atttcccacg	ctttgagttc	agatgggggg	atcatgtcta	cctgcggggc	gatgaagaaa	4620
acggtttccg	gggtaggggg	gatcagctgg	gaagaaagca	ggttcctgag	cagctgcgac	4680
ttaccgcagc	cggtggggcc	gtaaatcaca	cctattaccg	gctgcaactg	gtagttaaga	4740
gagctgcagc	tgccgtcatc	cctgagcagg	ggggccactt	cgtaagcat	gtccctgact	4800
cgcattgttt	ccctgaccaa	atccgccaga	aggcgctcgc	cgcccagcga	tagcagttct	4860
tgcaaggaag	caaagttttt	caacggtttg	agaccgtccg	ccgtaggcat	gcttttgagc	4920
gtttgaccaa	gcagttccag	gcgggtccac	agctcgggtca	cctgctctac	ggcatctcga	4980
tccagcatat	ctcctcgttt	cgcgggttgg	ggcggttttc	gctgtacggc	agtagtcgg	5040
gctcgtccag	acggggccagg	gtcatgtctt	tccacggggc	cagggtcctc	gtcagcgtag	5100
tctgggtcac	ggtgaagggg	tgcgtccgg	gctgcgcgct	ggccaggggtg	cgcttgaggc	5160
tggtcctgct	ggtgctgaag	cgctgccgg	cttcgccctg	cgcgtcggcc	aggtagcatt	5220
tgaccatggt	gtcatagtcc	agcccctccg	cggcggtggc	cttggcgcgc	agcttgccct	5280
tggaggaggc	gccgcacgag	gggcagtgca	gacttttgag	ggcgtagagc	ttgggcgcga	5340

gaaataccga ttccggggag taggcatccg cgccgcaggc cccgcagacg gtctcgcat 5400  
ccacgagcca ggtgagctct ggccgttcgg ggtcaaaaac caggtttccc ccatgctttt 5460  
tgatgcgttt cttacctctg gtttccatga gccggtgtcc acgctcgggtg acgaaaaggc 5520  
tgtccgtgtc cccgtataca gacttgagag gcctgtcctc gagcgggtgtt ccgcggtcct 5580  
cctcgtatag aaactcggac cactctgaga caaaggctcg cgtccaggcc agcacgaagg 5640  
aggctaagtg ggaggggtag cggtcgttgt ccactagggg gtccactcgc tccagggtgt 5700  
gaagacacat gtcgccctct tcggcatcaa ggaaggtgat tggttttag gtgtaggcca 5760  
cgtgaccggg tgttcctgaa ggggggctat aaaaggggtt gggggcgcgt tcgtcctcac 5820  
tctcttcgc atcgtgtct gcgagggcca gctgttgggg tgagtcgacg cgaggctgga 5880  
tggccttccc cattatgatt cttctcgtt ccggcggcat cgggatgcc gcgttgagg 5940  
ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaaggc cagcaaaagg 6000  
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggetccgc cccctgacg 6060  
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 6120  
accaggcgtt tccccctgga agctccctcg tgcgtctcc tgttcgacc ctgccgtta 6180  
ccgatacct gtccgccttt cctccctcgg gaagcgtggc gctttctcat agctcacgct 6240  
gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc 6300  
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa 6360  
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 6420  
taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag 6480  
tatttggtat ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt gtagctctt 6540  
gatccggcaa acaaacacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 6600  
cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 6660  
agtggaaaga aaactcacgt taagggattt tgggtcatgag attatcaaaa aggatcttca 6720  
cctagatcct tttaaattaa aaatgaagtt ttaaataaat cttaaagtata tatgagtaaa 6780  
cttggtctga cagttacca ttgcttaatca gtgaggcacc tatctcagcg atctgtctat 6840  
ttcgttcate catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 6900  
taccatctgg cccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt 6960  
tatcagcaat aaaccagcca gccggaagg cggagcgcag aagtggctct gcaactttat 7020  
ccgcctccat ccagtctatt aattgttgcc ggaagctag agtaagtagt tcgccagtta 7080  
atagtttgcg caacgttggt gccattgctg caggcatcgt ggtgtcacgc tcgtcgtttg 7140  
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt 7200  
tgtgcaaaaa agcggttagc tccttcggtc ctcgatcgt tgtcagaagt aagttggccg 7260  
cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 7320  
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 7380  
ggcgaccgag ttgctcttgc ccggcgtcaa cacgggataa taccgcgcca catagcagaa 7440  
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 7500  
cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt 7560  
ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 7620  
gaataagggc gacacggaaa tgttgaatac tcatactctt cttttttcaa tattattgaa 7680  
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 7740  
aacaatatag ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 7800  
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtcttcaag 7860  
aattggt 7867

&lt;210&gt; 43

&lt;211&gt; 7870

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 43

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60  
ggattgaagc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120  
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180  
tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgtg cgccgggtgtg cacaggaagt 240  
gacaattttc gcgcgggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300  
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360  
actcatagcg cgtaatatth gtctagggcc gcggggactt tgaccgttta cgtggagact 420  
cgcccagggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480  
tatagttagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540  
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600  
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660  
gagttccgcy ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720  
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780  
tgacgtcaat ggggtggagta tttacggtaa actgccact tggcagtaca tcaagtgtat 840  
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccg cgtggcattat 900  
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcac 960  
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020  
tcacggggat ttccaagtct ccaccccat gacgtcaatg ggagtttgtt ttggcaccaa 1080  
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggg 1140  
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200  
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260  
cgtcacgctc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320  
agaagtacca gcacctgtgg cgctggggct ggcgtgggg caccatgctg ctgggcatgc 1380  
tgatgatctg cagcgccacc gagaagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440  
ggaaggaggc caccaccacc ctgctctgcy ccagcgacgc caaggcctac gacaccgagg 1500  
tgcacaacgt gtgggcccacc cagcctgcy tgcaccacca ccccaacccc caggagggtg 1560  
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620  
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctggaag ctgaccccc 1680  
tgtgctgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740  
gccgcatgat catggagaag ggcgagatca agaactgcag ctctgctagc accagctgca 1800  
acaccagcgt gatcaccag gcctgcccc aaggtagctt cgagcccatc cccaaccact 1860  
actgcgccc cgccggcttc gccatcctga agtgcaagga caagaagttc aacggcaagg 1920  
gcccctgcac caacgtgagc accgtgcagt gcacccacgg catccgcccc gtggtgagca 1980  
cccagctgct gctgaacggg agcctggccg aggaggagg ggtgatccgc agcgtact 2040  
tcgcccgaac cgccaagggt atcatcgtgc agctgaacga gagcgtggag atcaactgca 2100  
cccgcacca caacaacacc cgcaagagca tccacatcgg ccccgccgc gccttctaca 2160  
ccaccggcga gatcatcggc gacatccgcc agggccactg caacctgagc cgcgccaagt 2220  
ggaacgacac cctgaacaag atcgtgatca agctgcgcga gcagttcggc aacaagacca 2280  
tcgtgttcaa gcacagcagc ggcggcgacc ccgagatcgt gacccacagc ttcaactgcg 2340  
gcggcgagtt cttctactgc aacagcgtc gcctgcccctg ccgcatcaag cagatcatca 2400  
acatgtggca gaagggtggc aaggccatgt acgcccccc catcagcggc cagatccgct 2460  
gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaacagc aacaacgaga 2520  
gcgagatctt ccgtccgggc ggcggcgaca tgcgcgacaa ctggcgagc gagctgtaca 2580  
agtacaagg ggtgaagatc gagcccttg gcggtggccc caccaaggcc aagcttaccg 2640  
tccaggccc ccagctgctg agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca 2700  
tcgaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc 2760  
gcaccctggc cgtggagcgc taccctgaagg accagcagca gctcgagcag atctggaacc 2820  
acaccacctg gatggagtgg gaccgcgaga tcaacaacta caccagcctg atccacagcc 2880  
tgatcgagga gagccagaac cagcaggaga agaacgagca ggagctgctg gagctggaca 2940  
agtgggccag cctgtggaac tggttcaaca tcaccaactg gctgtggtga ggatccagat 3000  
ctgctgtgcc ttctagtgtc cagccatctg ttgtttgccc ctccccgctg ctttccttga 3060  
ccctggaagg tgcactccc actgtcctt cctaataaaa tgaggaaatt gcacgcatt 3120  
gtctgagtag gtgtcattct attctggggg gtgggggtgg gcagcacagc aagggggagg 3180  
attgggaaga caatagcagg catgctgggg atgcgggtgg ctctatgggt acccagggcc 3240  
gcataacttc gtataatgta tgcatacga agttataaga tctgtactga aatgtgtggg 3300



cgtggcttaa ggggtgggaaa gaatatataa ggtggggggtc ttatgtagtt ttgtatctgt 3360  
tttgcagcag ccgccgccgc catgagcacc aactcgtttg atggaagcat tgtgagctca 3420  
tatttgacaa cgcgcgatgcc cccatgggcc ggggtgctgc agaatgtgat gggctccagc 3480  
attgatggtc gcccgcctct gcccgcaaac tctactacct tgacctacga gaccgtgtct 3540  
ggaacgccgt tggagactgc agcctccgcc gccgcttcag ccgctgcagc caccgcccgc 3600  
gggattgtga ctgactttgc ttctctgagc ccgcttgcaa gcagtgcagc ttcccgttca 3660  
tccgcccgcg atgacaagtt gacggctctt ttggcacaat tggattcttt gaccggggaa 3720  
cttaatgtcg tttctcagca gctgttgat ctgcgccagc aggtttctgc cctgaaggct 3780  
tcctcccctc ccaatgcggt ttaaaacata aataaaaaac cagactctgt ttggatttgg 3840  
atcaagcaag tgtcttgctg tctttattta ggggttttgc gcgcgcggtg ggcccgggac 3900  
cagcggctct gccggttgag ggtcctgtgt attttttcca ggacgtggtg aaggtgactc 3960  
tggatgttca gatacatggg cataagcccg tctctggggt ggaggtagca ccactgcaga 4020  
gcttcatgct gcgggggtgt gttgtagatg atccagtcgt agcaggagcg ctgggctgtg 4080  
tgcctaaaaa tgtctttcag tagcaagctg attgccaggg gcaggccctt ggtgtaagtg 4140  
tttacaagc ggtaagctg ggatgggtgc atacgtgggg atatgagatg catcttggac 4200  
tgtattttta ggttggtatg gttcccagcc atatccctcc ggggattcat gttgtgcaga 4260  
accaccagca cagtgtatcc ggtgcacttg ggaaatttgt catgtagctt agaaggaaat 4320  
gcgtggaaga acttgagagc gcccttgtga cctccaagat tttccatgca ttcgtccata 4380  
atgatggcaa tgggcccacg ggccggcgcc tgggcgaaga tatttctggg atcactaacg 4440  
tcatagtgtg gttccaggat gagatcgtca taggccattt ttacaaagcg cgggcggagg 4500  
gtgccagact gcggtataat ggttccatcc ggcccagggg cgtagttacc ctccagatt 4560  
tgcatttccc acgctttgag ttcagatggg gggatcatgt ctacctgcgg ggcgatgaag 4620  
aaaacggttt ccggggtagg ggagatcagc tgggaagaaa gcaggttcct gagcagctgc 4680  
gacttaccgc agccgggtgg cccgtaaatc acacctatta ccggctgcaa ctggtagtta 4740  
agagagctgc agctgccgtc atccctgagc agggggggcca ctctgttaag catgtccctg 4800  
actcgcagt tttccctgac caaatccgcc agaaggcgct cgcgcgccag cgatagcagt 4860  
tcttgcaagg aagcaaagtt tttcaacggt ttgagaccgt ccgcgtagg catgttttg 4920  
agcgtttgac caagcagttc caggcggtcc cacagctcgg tcacctgctc tacggcatct 4980  
cgatccagca tatctcctcg tttcgcgggt tggggcggtc ttcgctgtac ggcagtagtc 5040  
ggtgctcgtc cagacggggc agggctcatgt ctttccacgg gcgcagggtc ctgctcagcg 5100  
tagtctgggt cacggtgaag ggggtgctgc cgggctgcgc gctggccagg gtgcgcttga 5160  
ggctggctct gctgggtgctg aagcgtgcc ggtcttcgcc ctgcgcgtcg gccaggtagc 5220  
atltgaccat ggtgtcatag tccagcccct ccgcggcggt gcccttggcg cgcagcttgc 5280  
ccttgaggga ggcgccgcac gaggggcagt gcagactttt gagggcgtag agcttgggcg 5340  
cgagaaatac cgattccggg gagtaggcat ccgcgccgca ggcccgcag acggtctcgc 5400  
attccacgag ccaggtgagc tctggccggt cggggtcaaa aaccaggttt ccccatgct 5460  
ttttgatgct tttcttacct ctggtttcca tgagccggtg tccacgctcg gtgacgaaaa 5520  
ggctgtccgt gtccccgtat acagacttga gaggcctgtc ctcgagcgggt gttccgcggt 5580  
cctcctcgta tagaaactcg gaccactctg agacaaaggc tcgctccag gccagcacga 5640  
aggaggctaa gtgggagggg tagcggctgt tgtccactag ggggtccact cgctccaggg 5700  
tgtgaagaca catgtcgccc tcttcggcat caaggaaggt gattggtttg taggtgtagg 5760  
ccacgtgacc ggggtgttct gaaggggggc tataaaaggg ggtgggggcg cgttcgtcct 5820  
cactctcttc cgcctcgtg tctgcgagg ccagctgttg ggggtgagtc acgcgaggct 5880  
ggatggcctt ccccatatg attcttctcg ctcccgcg catcgggatg cccgcgttgc 5940  
aggccatgct gtccaggcag gtagatgacg accatcaggg acagcttcaa ggccagcaaa 6000  
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 6060  
acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaaccgcaca ggactataaa 6120  
gataccaggc gtttccccct ggaagctccc tcgtgcgtc tcctgttccg accctgccgc 6180  
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 6240  
gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 6300  
ccccggttca gcccgaccgc tgcgccttat ccggttaacta tcgtcttgag tccaaccggg 6360  
taagacacga cttatcgcca ctggcagcag cactggtaaa caggattagc agagcgagg 6420  
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 6480  
cagtatttgg tatctgcgt ctgctgaagc cagttacctt cggaaaaaga gttggtagct 6540  
cttgatccgg caaacaacc accgctggta gcggtgggtt ttttgtttgc aagcagcaga 6600

```

ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg 6660
ctcagtggaa cgaaaactca cgttaagggg ttttggtcat gagattatca aaaaggatct 6720
tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 6780
aaacttggtc tgacagttac caatgcttaa tcagtgaggg acctatctca gcgatctgtc 6840
tatttcgttc atccatagtt gctgactcc cgtcgtgta gataactacg atacgggagg 6900
gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag 6960
atztatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc cctgcaactt 7020
tatccgcctc catccagtct attaatgttt gccgggaagc tagagtaagt agttcgccag 7080
ttaatagttt gcgcaacggt gttgccattg ctgcaggcat cgtgggtgtc cgctcgtcgt 7140
ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccc 7200
tggttgtaa aaaagcgggt agctccttcg gtcctccgat cgttggtcaga agtaagttgg 7260
ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 7320
ccgtaagatg cttttctgtg actgggtgag actcaaccaa gtcattctga gaatagtgt 7380
tgccggcgacc gagttgctct tgcgcggcgt caacacggga taataccgcg ccacatagca 7440
gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 7500
taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 7560
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 7620
agggaaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 7680
gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 7740
ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 7800
ccattattat catgacatta acctataaaa ataggcgtat caccaggccc tttcgtcttc 7860
aagaattggt                                     7870

```

&lt;210&gt; 44

&lt;211&gt; 7885

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 44

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttggt acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgatgatg tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaata ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaataatt gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagtggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta ttacggtaa actgccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgcccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960
gctattacca tggatgatgc gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccaccccat gacgtcaatg ggagtttgtt ttggcacc 1080
aatcaacggg actttccaaa atgtcgtaac aactccgcc cattgacgca aatgggcgg 1140
aggcgtgtac ggtgggagggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tgagagacgc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260

```

```

cgtcaccgtc gtcgacacgt gtgatcagat atcgcgggccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgccacc gagaagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtgggcccacc cagcctgctg tggcccaccga ccccaacccc caggagggtg 1560
tgctgggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctgtaag ctgaccccc 1680
tgtgctgag cctgaagtgc accgacctga agaacgacac caacaccaac agcagcagcg 1740
gccgcatgat catggagaag ggcgagatca agaactgcag cttcaacatc agcaccagca 1800
tccgcgga ggtgcagaag gaggtagcct tcttctacaa gctggacatc atccccatcg 1860
acaacgacac caccagctac agcctgacca gctgcaacac cagcgtgatc acccaggcct 1920
gccccaaagt gagcttcgag cccatcccca accactactg cgcccccgcc ggcttcgcca 1980
tcctgaagtg caaggacaag aagttcaacg gcaaggggcc ctgcaccaac gtgagcaccg 2040
tgcagtgcac ccacggcatc cgccccgtgg tgagcaccga gctgctgctg aacggtagcc 2100
tggccgagga ggagggtggtg atccgcagcg ctaacttcgc cgacaacgcc aaggtagtca 2160
tcgtgcagct gaacgagagc gtggagatca actgcacccg cgctagcgcc cactgcaacc 2220
tgagccgcg caagtggaa gacaccctga acaagatcgt gatcaagctg cgcgagcagt 2280
tcggcaacaa gaccatcgtg ttcaagcaca gcagcgggcg cgaccccgag atcgtgaccc 2340
acagcttcaa ctgcgggcggc gaggttcttct actgcaacag cgctagcctg ccctgccgca 2400
tcaagcagat catcaacatg tggcagaagg tgggcaaggc catgtacgcc cccccatca 2460
gcgccagat ccgctgcagc agcaacatca ccggcctgct gctgacccgc gacggcgcca 2520
acagcaacaa cgagagcgag atcttccgtc cgggcgggcg cgacatgcgc gacaactggc 2580
gcagcgagct gtacaagtac aagggtggtga agatcgagcc cctgggctg gccccacca 2640
aggccaagct taccgtccag gcccgccagc tgctgagcgg catcgtgcag cagcagaaca 2700
acctgctgcg cgccatcgag gcccgagcgc acctgctgca gctgaccgtg tggggcatca 2760
agcagctgca ggcccgacc ctggccgtgg agcgctacct gaaggaccag cagcagctcg 2820
agcagatctg gaaccacacc acctggatgg agtgggaccg cgagatcaac aactacacca 2880
gcctgatcca cagcctgatc gaggagagcc agaaccagca ggagaagaac gagcaggagc 2940
tgctggagct ggacaagtgg gccagcctgt ggaactggtt caacatcacc aactggctgt 3000
gggtgaggatc cagatctgct gtgccttcta gttgccagcc atctgttgtt tgccccctcc 3060
ccgtgccttc cttgaccctg gaaggtgcca ctccactgt cctttcctaa taaaatgagg 3120
aaattgcate gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagc 3180
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 3240
tgggtaccca gggccgcata acttcgtata atgtatgcta tacgaagtta taagatctgt 3300
actgaaatgt gtgggcgtgg cttaagggtg ggaaagaata tataagggtg ggggtcttatg 3360
tagttttgta tctgttttgc agcagccgcc gccgccatga gcaccaactc gtttgatbga 3420
agcattgtga gctcatatct gacaacgcgc atgcccccat gggccggggg gcgtcagaat 3480
gtgatgggct ccagcattga tggtcgcccc gtccctgcccg caaactctac taccttgacc 3540
tacgagaccg tgtctggaac gccgttgag actgcagcct ccgccgccgc ttcagccgct 3600
gcagccaccg cccgcgggat tgtgactgac tttgctttcc tgagcccgct tgcaagcagt 3660
gcagcttccc gttcatccgc ccgcgatgac aagttgacgg ctcttttggc acaattggat 3720
tctttgaccc gggaaacttaa tgtcgtttct cagcagctgt tggatctgcg ccagcagggt 3780
tctgccctga aggttcttc cctcccaat gcggtttaa acataaataa aaaaccagac 3840
tctgtttgga tttggatcaa gcaagtgtct tgctgtcttt atttaggggt tttgcgcgcg 3900
cggtaggccc gggaccagcg gtctcggtcg ttgagggtcc tgtgtatttt ttccaggacg 3960
tggtaaagg gactctggat gttcagatac atgggcataa gccgctctct ggggtggagg 4020
tagcaccact gcagagcttc atgctgcggg gtggtgttgt agatgatcca gtcgtagcag 4080
gagcgctggg cgtgggtgcct aaaaatgtct ttcagtagca agctgattgc caggggcagg 4140
cccttggtgt aagtgtttac aaagcgggtta agctgggatg ggtgcatacg tggggatatg 4200
agatgcatct tggactgtat ttttaggttg gctatgttcc cagccatata cctccgggga 4260
ttcatgttgt gcagaaccac cagcacagt tatccggtgc acttgggaaa tttgtcatgt 4320
agcttagaag gaaatgcgtg gaagaacttg gagacgccct tgtgacctcc aagattttcc 4380
atgcattcgt ccataatgat ggcaatgggc ccacggggcg cgccctgggc gaagatatct 4440
ctgggatcac taacgtcata gttgtgttcc aggatgagat cgtcataggc catttttaca 4500
aagcgcgggc ggagggtgcc agactgcggt ataatgggtc catccggccc aggggcgtag 4560

```



```

ttaccctcac agatttgcac ttcccacgct ttgagttcag atgggggggat catgtctacc 4620
tgcggggcgga tgaagaaaac ggtttccggg gtaggggaga tcagctggga agaaagcagg 4680
ttcctgagca gctgcgactt accgcagccg gtggggccgt aaatcacacc tattaccggc 4740
tgcaactggt agttaagaga gctgcagctg ccgtcatccc tgagcagggg ggccacttcg 4800
ttaagcatgt ccttgactcg catgttttcc ctgaccaaab ccgccagaag gcgctcgccg 4860
cccagcgata gcagttcttg caaggaagca aagtttttca acggtttgag accgtccgcc 4920
gtaggcatgc ttttgagcgt ttgaccaagc agttccaggc ggtcccacag ctcggtcacc 4980
tgctctacgg catctcgatc cagcatatct cctcgtttcg cgggttgggg cggcttttcg 5040
tgtacggcag tagtcggtgc tcgtccagac gggccagggt catgtcttcc cagggcgca 5100
gggtcctcgt cagcgtagtc tgggtcacgg tgaaggggtg cgctccgggc tgcgcgctgg 5160
ccagggtgcg cttgaggctg gtctgctgg tgctgaagcg ctgccggtct tcgccctgcg 5220
cgctcgccag gtagcatttg accatggtgt catagtccag cccctccgcg gcgtggccct 5280
tggcgcgag cttgcccttg gaggaggcgc cgcacgaggg gcagtcgaga cttttgaggg 5340
cgtagagctt gggcgcgaga aataccgatt ccggggagta ggcacccgcg ccgcaggccc 5400
cgcagacggt ctgcattcc acgagccagg tgagctctgg ccgttcgggg tcaaaaacca 5460
ggtttccccc atgctttttg atgcgtttct tacctctggt ttccatgagc cgggtgtccac 5520
gctcggtgac gaaaaggctg tccgtgtccc cgtatacaga cttgagaggc ctgtcctcga 5580
gcggtgttcc gcggtcctcc tcgtatagaa actcggacca ctctgagaca aaggctcgcg 5640
tccaggccag cacgaaggag gctaagtggg aggggtagcg gtcgttgtcc actagggggg 5700
ccactcgctc cagggtgtga agacacatgt cgcctcttcc ggcacaaagg aagggtgattg 5760
gtttgtaggt gtaggccacg tgaccgggtg ttctgaagg ggggtataa aaggggtgg 5820
gggcgcgttc gtctcactc tcttcgcgt cgtgtctgc gagggccagc tgttgggggtg 5880
agtcgacgcg aggttggtg gccttcccca ttatgattct tctcgcttcc ggccgcatcg 5940
ggatgcccgc gttgcaggcc atgctgtcca ggcaggtaga tgacgacct cagggacagc 6000
ttcaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 6060
ggctccgccc cctgacgag catcacaana atcgacgctc aagtcagagg tggcgaaacc 6120
cgacaggact ataaagatac caggcgttcc cccctggaag ctccctcggt cgctctcctg 6180
ttccgacctc gccgttacc ggatacctgt ccgccttctc ccttcggga agcgtggcg 6240
tttctcatag ctacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg 6300
gctgtgtgca cgaaccccc gttagcccg accgctgcgc cttatccgtt aactatcgtc 6360
ttgagtccaa cccggttaaga cagacttat cgcactggc agcagccact ggtaacagga 6420
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg 6480
gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 6540
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg 6600
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 6660
ctacggggtc tgacgctcag tggaaagaaa actcacgtta agggattttg gtcattgagat 6720
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 6780
aaagtatata tgagtaaaact tgggtctgaca gttaccaatg cttaatcagt gaggcacct 6840
tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa 6900
ctacgatacg ggagggttta ccatctggcc ccagtgtgc aatgataacc cgagaccac 6960
gctcaccggc tccagattta tcagcaataa accagccagc cggaaggggc gagcgcagaa 7020
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttggtgcccg gaagctagag 7080
taagtagttc gccagttaat agtttgcgca acgttggtgc cattgctgca ggcacgtgg 7140
tgtcacgctc gtcgttttgg atggcttcat tcagctccgg ttcccaacga tcaaggcgag 7200
ttacatgac cccatgttg tgcaaaaaag cggttagctc cttcggctct ccgacgttg 7260
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc 7320
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcac 7380
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaaca cgggataata 7440
ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa 7500
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact cgtgcacca 7560
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc 7620
aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc atactcttcc 7680
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 7740
aatgtattta gaaaaataaa caaatagggt ttccgcgcac atttccccga aaagtgcac 7800
ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga 7860

```

ggcccttttcg tcttcaagaa ttggt

7885

&lt;210&gt; 45

&lt;211&gt; 7789

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 45

```
ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggt ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcaactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcate 960
gctattacca tggatgatgc gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccaccccat gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcacccgtc gtcgacacgt gtgatcagat atcgcgggcc ctctagaatg cgcgtaagg 1320
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgccacc gagaagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcacaacgt gtgggccacc cagcctgcg tgcaccacca cccaacccc caggagggtg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgctgtaag ctgaccccc 1680
tgtgctgag cctgaagtgc accgacgcta gcaccagctg caacaccagc gtgatcacc 1740
aggcctgccc caaggtgagc ttcgagccca tccccacca ctactgcgc cccgccggct 1800
tcgccatcct gaagtgcaag gacaagaagt tcaacggcaa gggccctgc accaacgtga 1860
gcaccgtgca gtgcaccac ggcatccgcc ccgtggtgag caccagctg ctgctgaacg 1920
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgccgac aacgccaagg 1980
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcgct agcgcacct 2040
gcaacctgag ccgcgccaag tggaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2100
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2160
tgaccacag cttcaactgc ggcgcgagt tcttctactg caacagcacc cagctgttca 2220
acagcacctg gttcaacagc acctggagca ccgagggcag caacaacacc gagggcagcg 2280
acaccatcac cctgccctgc cgcacaaagc agatcatcaa catgtggcag aagggtgggca 2340
aggccatgta cgccccccc atcagcggcc agatccgctg cagcagcaac atcaccggcc 2400
tgctgctgac ccgcgacggc ggcaacagca acaacgagag cgagatcttc cgtccgggcg 2460
gcggcgacat gcgcgacaac tggcgacgag agctgtacaa gtacaagggtg gtgaagatcg 2520
```

agccccctggg	cgtgggcccc	accaaggcca	agcttaccgt	ccaggccccg	cagctgctga	2580
gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggccccag	cagcacctgc	2640
tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccc	caccctggcc	gtggagcgct	2700
acctgaagga	ccagcagcag	ctcgagcaga	tctggaacca	caccacctgg	atggagtggg	2760
accgcgagat	caacaactac	accagcctga	tccacagcct	gatcgaggag	agccagaacc	2820
agcaggagaa	gaacgagcag	gagctgctgg	agctggacaa	gtgggccagc	ctgtggaact	2880
ggttcaacat	caccaactgg	ctgtggtgag	gatccagatc	tgtgtgcct	tctagtgtcc	2940
agccatctgt	tgtttgcccc	tccccctgtc	cttccttgac	cctggaaggt	gccactccca	3000
ctgtcctttc	ctaataaaat	gaggaaattg	catcgcatgt	tctgagtagg	tgtcattcta	3060
ttctgggggg	tgggggtggg	cagcacagca	agggggagga	ttgggaagac	aatagcaggc	3120
atgctgggga	tgcggtgggc	tctatgggta	cccaggggcc	cataacttcg	tataatgtat	3180
gctatacgaa	gttataagat	ctgtactgaa	atgtgtgggc	gtggcttaag	ggtgggaaag	3240
aatatataag	gtgggggtct	tatgtagttt	tgtatctgtt	ttgcagcagc	cgccgccgcc	3300
atgagcacca	actcgtttga	tggaaagcatt	gtgagctcat	atltgacaac	gcgcagtccc	3360
ccatggggcc	gggtgcgtca	gaatgtgatg	ggctccagca	ttgatgggtc	ccccgtcctg	3420
cccgcaaact	ctactacctt	gacctacgag	accgtgtctg	gaacgccgtt	ggagactgca	3480
gcctccgccg	ccgcttcagc	cgctgcagcc	accgccccgc	ggattgtgac	tgactttgct	3540
ttcctgagcc	cgcttgcaag	tagtgcagct	tcccgttcat	ccgccccgca	tgacaagttg	3600
acggctcttt	tggcacaatt	ggattctttg	accggggaac	ttaatgtcgt	ttctcagcag	3660
ctgttggatc	tgcgccagca	ggtttctgcc	ctgaaggctt	cctccccctc	caatgcgggt	3720
taaaacataa	ataaaaaacc	agactctggt	tggatttgga	tcaagcaagt	gtcttgctgt	3780
ctttatbttag	gggttttgcg	cgcgcggtag	gcccgggacc	agcggctctc	gtcgttgagg	3840
gtcctgtgta	ttttttccag	gacgtggtaa	aggtgactct	ggatgttcag	atacatgggc	3900
ataagcccgt	ctctgggggt	gaggtagcac	cactgcagag	cttcatgctg	cggggtgggt	3960
ttgtagatga	tccagtcgta	gcaggagcgc	tgggcgtggt	gcctaaaaat	gtctttcagt	4020
agcaagctga	ttgccagggg	caggcccttg	gtgtaagtgt	ttacaaagcg	gttaagctgg	4080
gatgggtgca	tacgtgggga	tatgagatgc	atcttggtact	gtatttttag	gttggctatg	4140
ttcccagcca	tatccctccg	gggattcatg	ttgtgcagaa	ccaccagcac	agtgtatccg	4200
gtgcacttgg	gaaatttgte	atgtagctta	gaaggaaatg	cgtggaagaa	cttgagagac	4260
cccttgtgac	ctccaagatt	ttccatgcat	tgcgtccata	tgatggcaat	gggcccacgg	4320
gcggcggcct	gggcgaagat	atltctggga	tcactaacgt	catagtgtgt	ttccaggatg	4380
agatcgtcat	aggccatttt	tacaaagcgc	gggcggaggg	tgccagactg	cggtataatg	4440
gttccatccg	gcccaggggc	gtagttaccc	tcacagattt	gcatttccca	cgctttgagt	4500
tcagatgggg	ggatcatgtc	tacctgcggg	gcgatgaaga	aaacgggttc	cggggtaggg	4560
gagatcagct	gggaagaaag	cagggtccctg	agcagctgcg	acttaccgca	gccgggtggc	4620
ccgtaaatca	cacctattac	cggttgcaac	tggtagttaa	gagagctgca	gctgccgtca	4680
tccctgagca	ggggggccac	ttcgttaagc	atgtccctga	ctcgcattgt	ttccctgacc	4740
aaatccgcca	gaaggcgctc	gccgccagc	gatagcagtt	cttgcaagga	agcaaagttt	4800
ttcaacgggt	tgagaccgtc	cgcgtaggc	atgcttttga	gcgtttgacc	aagcagttcc	4860
aggcggtccc	acagctcggt	cacctgctct	acggcatctc	gatccagcat	atctcctcgt	4920
ttcgcgggtt	ggggcggtt	tcgctgtacg	gcagtagtcg	gtgctcgtcc	agacgggcca	4980
gggtcatgtc	tttccacggg	cgcagggtcc	tcgtcagcgt	agtctgggtc	acgggtgaagg	5040
gggtgcgtcc	gggtgcgcg	ctggccaggg	tgcgcttgag	gctggctcctg	ctggtgctga	5100
agcgtgccc	gtcttcgccc	tgcgcgtcgg	ccaggtagca	tttgaccatg	gtgtcatagt	5160
ccagccccctc	cgcggcgtgg	cccttgggcg	gcagcttgcc	cttgaggagag	gcgccgcacg	5220
aggggcagtg	cagacttttg	agggcgtaga	gcttggggcg	gagaaatacc	gattccgggg	5280
agtaggcata	cgcgccgcag	gccccgcaga	cggctctcgca	ttccacgagc	cagggtgagct	5340
ctggccgttc	gggtcaaaa	accaggtttc	ccccatgctt	tttgatgcgt	ttcttacctc	5400
tggtttccat	gagccggtgt	ccacgctcgg	tgacgaaaag	gctgtccgtg	tccccgtata	5460
cagacttgag	aggcctgtcc	tcgagcgggt	ttccgcgggt	ctcctcgtat	agaaactcgg	5520
accactctga	gacaaaggct	cgcgtccagg	ccagcacgaa	ggaggctaag	tgggaggggt	5580
agcggtcgtt	gtccactagg	gggtccactc	gctccagggg	gtgaagacac	atgtcgccct	5640
cttcggcatc	aaggaagggt	attggtttgt	aggtgtaggc	cacgtgaccg	ggtgttcctg	5700
aaggggggct	ataaaagggg	gtggggggcg	gttcgtcctc	actctcttcc	gcatacgtgt	5760
ctgcgagggc	cagctgttgg	ggtgagtcga	cgcgaggctg	gatggccttc	cccattatga	5820

```

ttcttctcgc ttccggcggc atcgggatgc ccgcgttgca ggccatgctg tccaggcagg 5880
tagatgacga ccatcaggga cagcttcaag gccagcaaaa ggccagggaac cgtaaaaagg 5940
ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 6000
gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 6060
gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 6120
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctgagttcgg 6180
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 6240
gcgccttata cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 6300
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt 6360
tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatattgg atctgcgctc 6420
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 6480
ccgctggtag cgggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 6540
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaaac gaaaactcac 6600
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 6660
aaaaatgaag ttttaaatac atctaaagta tatatgagta aacttgggtc gacagttacc 6720
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 6780
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 6840
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 6900
cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta 6960
ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 7020
ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tgggatggct tcattcagct 7080
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 7140
gctccttcgg tcctccgata gttgtcagaa gtaagttggc cgcagtgtta tcaactatgg 7200
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 7260
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 7320
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7380
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 7440
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 7500
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 7560
aatgttgaat actcatactc ttcttttttc aatattattg aagcatttat cagggttatt 7620
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7680
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7740
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattgtt 7789

```

&lt;210&gt; 46

&lt;211&gt; 7792

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 46

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgt cgccggtgta cacaggaagt 240
gacaattttc gcgcgggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccagggt tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta rattggctca tgtccaacat taccgccatg ttgacattga 600

```



ttattgacta	gttattaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	660
gagttccgcg	ttacataact	tacggtaaatt	ggccccgcctg	gctgaccgcc	caacgacccc	720
cgccccattga	cgtcaataat	gacgtatggt	cccatagtaa	cgccaatagg	gactttccat	780
tgacgtcaat	gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	840
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggccccg	ctggcattat	900
gcccagtaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcatc	960
gctattacca	tgggtgatgcg	gtttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	1020
tcacggggat	ttccaagtct	ccacccatt	gacgtcaatg	ggagtttggt	ttggcaccaa	1080
aatcaacggg	actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggcggt	1140
aggcgtgtac	ggtgggaggt	ctatataagc	agagctcggt	tagtgaaccg	tcagatcgcc	1200
tggagacgcc	atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	1260
cgtcaccgtc	gtcgacacgt	gtgatcagat	atcgcgggccg	ctctagaatg	cgcggtgaagg	1320
agaagtacca	gcacctgtgg	cgctggggct	ggcgctgggg	caccatgctg	ctgggcatgc	1380
tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgcacaacgt	gtgggccacc	cacgcctgcg	tgcccaccga	ccccaacccc	caggagggtg	1560
tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgctggaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacgcta	gcaccagctg	caacaccagc	gtgatcacc	1740
aggcctgccc	caaggtgagc	ttcgagccca	tccccaacca	ctactgcgcc	cccgcgggt	1800
tcgccatcct	gaagtgcaag	gacaagaagt	tcaacggcaa	gggcccctgc	accaacgtga	1860
gcaccgtgca	gtgcaccac	ggcatccgcc	ccgtggtgag	caccagctg	ctgctgaacg	1920
gtagcctggc	cgaggaggag	gtggtgatcc	gcagcgctaa	cttcgcccgc	aacgccaaag	1980
tgatcatcgt	gcagctgaac	gagagcgtgg	agatcaactg	caccgcccc	aacaacaaca	2040
cccgaagag	catccacatc	ggccccggcc	gcgccttcta	caccaccggc	gagatcatcg	2100
gcgacatccg	ccaggccac	tgcaacctga	gccgcgccaa	gtggaacgac	accctgaaca	2160
agatcgtgat	caagctgcgc	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2220
gcggcggcga	ccccgagatc	gtgaccacac	gcttcaactg	cggcggcgag	ttcttctact	2280
gcaacagcgc	tagcctgccc	tgccgcata	agcagatcat	caacatgtgg	cagaagggtg	2340
gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	ctgcagcagc	aacatcaccg	2400
gcctgctgct	gaccgcgcgac	ggcggaaca	gcaacaacga	gagcgagatc	ttccgtccgg	2460
gcggcggcga	catgcgcgac	aactggcgca	gcgagctgta	caagtacaag	gtggtgaaga	2520
tcgagccct	gggctggcc	cccaccaagg	ccaagcttac	cgtccaggcc	cgccagctgc	2580
tgagcggcat	cgtgcagcag	cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	2640
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcaggc	ccgcaccctg	gccgtggagc	2700
gctacctgaa	ggaccagcag	cagctcgagc	agatctggaa	ccacaccacc	tggatggagt	2760
gggaccgcga	gatcaacaac	tacaccagcc	tgatccacag	cctgatcgag	gagagccaga	2820
accagcagga	gaagaacgag	caggagctgc	tggagctgga	caagtgggcc	agcctgtgga	2880
actggttcaa	catcaccaac	tggctgtggt	gaggatccag	atctgctgtg	ccttctagtt	2940
gccagccatc	tgttgtttgc	ccctcccccg	tgcttctctt	gaccctggaa	ggtgccactc	3000
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3060
ctattctggg	gggtgggggtg	gggcagcaca	gcaaggggga	ggattgggaa	gacaatagca	3120
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccagggg	ccgcataact	tcgtataatg	3180
tatgctatac	gaagtataaa	gatctgtact	gaaatgtgtg	ggcgtggctt	aagggtggga	3240
aagaatatat	aagggtggggg	tcttatgtag	ttttgtatct	gttttgacgc	agccgccgcc	3300
gccatgagca	ccaactcggt	tgatggaagc	attgtgagct	catatttgac	aacgcgcagt	3360
cccccatggg	ccgggggtgcg	tcagaatgtg	atgggctcca	gcattgatgg	tcgccccgtc	3420
ctgcccga	actctactac	cttgacctac	gagaccgtgt	ctggaacgcc	gttggagact	3480
gcagcctccg	ccgcgccttc	agccgctgca	gccaccgccc	gcgggattgt	gactgacttt	3540
gctttcctga	gcccgccttg	aagcagtgca	gcttcccgtt	catccgcccg	cgatgacaag	3600
ttgacggctc	ttttggcaca	attggattct	ttgaccgggg	aacttaattg	cgtttctcag	3660
cagctggttg	atctgcgcca	gcaggtttct	gccctgaagg	cttctcccc	tcccaatgcg	3720
gtttaaaaca	taaataaaaa	accagactct	gtttggattt	ggatcaagca	agtgtcttgc	3780
tgtctttatt	taggggtttt	gcgcgcgcgg	taggcccggg	accagcggtc	tcggtcggtg	3840
agggtcctgt	gtattttttc	caggacgtgg	taaagggtgac	tctggatgtt	cagatacatg	3900

ggcataagcc cgtctctggg gtggaggtag caccactgca gagcttcatg ctgcggggtg 3960  
gtgttgtaga tgatccagtc gtagcaggag cgtgggcgt ggtgcctaaa aatgtcttct 4020  
agtagcaagc tgattgccag gggcaggccc ttggtgtaag tgtttacaaa gcggttaagc 4080  
tgggatgggt gcatacgtgg ggatatgaga tgcattcttg actgtatttt taggttggt 4140  
atgttcccag ccatatccct ccggggatct atgttgtgca gaaccaccag cacagtgtat 4200  
ccggtgcact tgggaaattt gtcattgtgc ttagaaggaa atgcgtggaa gaacttgag 4260  
acgcccttgt gacctccaag attttccatg cattcgtcca taatgatggc aatgggcca 4320  
cgggcggcgg cctgggcgaa gatatttctg ggatcactaa cgtcatagtt gtgttccagg 4380  
atgagatcgt cataggccat ttttacaaaag cgcgggcgga ggggtgccaga ctgcggtata 4440  
atggttccat ccggcccagg ggcgtagtta ccctcacaga tttgcatttc ccacgctttg 4500  
agttcagatg gggggatcat gtctacctgc ggggcgatga agaaaacggc ttccggggta 4560  
ggggagatca gctgggaaga aagcaggttc ctgagcagct gcgacttacc gcagccggtg 4620  
ggcccgtaaa tcacacctat taccggctgc aactggtagt taagagagct gcagctgccg 4680  
tcatccctga gcaggggggc cacttcgtta agcatgtccc tgactcgcct gtttccctg 4740  
accaaattcc ccagaaggcg ctgcgccccc agcgatagca gttcttgcaa ggaagcaaag 4800  
tttttcaacg gtttgagacc gtccgcgcta ggcattgttt tgagcgtttg accaagcagt 4860  
tccaggcggc ccacagctc ggtcacctgc tctacggcat ctcgatccag catatctcct 4920  
cgtttcgcgg gttggggcgg ctttcgctgt acggcagtag tgggtgctcg tccagacggg 4980  
ccagggtcat gtcttccac gggcgcaggg tctcgtcag cgtagtctgg gtcacggtga 5040  
aggggtgcgc tccgggctgc gcgctggcca ggggtgcgct gaggctggtc ctgctgggtg 5100  
tgaagcgtg ccggtcttcg cctgcgcgt cggccaggta gcatttgacc atggtgtcat 5160  
agtcagccc ctccgcggcg tggcccttgg cgcgcagctt gcccttggag gaggcgcgc 5220  
acgaggggca gtgcagactt ttgagggcgt agagcttggg cgcgagaaat accgattccg 5280  
gggagtaggc atccgcgccg caggccccgc agacggtctc gcattccacg agccagggtga 5340  
gctctggccg ttcgggggtca aaaaccaggc tcccccatg ctttttgatg cgtttcttac 5400  
ctctggtttc catgagccgg tgtccacgct cggtgacgaa aaggctgtcc gtgtccccgt 5460  
atacagactt gagaggcctg tctcgcagcg gtgttccgcg gtcctcctcg tatagaaact 5520  
cggaccactc tgagacaaag gctcgcgtcc aggcagcac gaaggaggct aagtgggagg 5580  
ggtagcggtc gttgtccact aggggggtcca ctgcctccag ggtgtgaaga cacatgtcgc 5640  
cctcttcggc atcaagggaag gtgattgggt tgtaggtgta ggccacgtga ccgggtgttc 5700  
ctgaaggggg gctataaaag ggggtggggg cgcgttcgtc ctactctct tccgcacgc 5760  
tgtctgcgag ggccagctgt tggggtgagt cgacgcgagg ctggatggcc tccccatta 5820  
tgattcttct cgttccggc ggcacggga tgccgcggt gcaggccatg ctgtccaggc 5880  
aggtagatga cgaccatcag ggacagcttc aaggccagca aaaggccagg aaccgtaaaa 5940  
agggcgcgtt gctggcgttt tcccataggc tccgcccccc tgacgagcat cacaaaaatc 6000  
gacgtcaag tcagagggtg cgaaaccgca caggactata aagataccag gcgtttcccc 6060  
ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 6120  
cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 6180  
cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc 6240  
gctgcgcctt atccggtaac tctcgtcttg agtccaacct ggtaagacac gacttatcgc 6300  
cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 6360  
agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg 6420  
ctctgctgaa gccagttacc ttcggaaaaa gagtggtag ctcttgatcc ggcaaaaaa 6480  
ccaccgctgg tagcgggtgt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 6540  
gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 6600  
cacgttaagg gattttggtc atgagattat caaaaaggat ctccacctag atccttttaa 6660  
attaaaaatg aagtttttaa tcaatctaaa gtatatatga gtaaaacttg tctgacagtt 6720  
accaatgctt aatcagttag gcacctatct cagcgatctg tctatttctg tcatccatag 6780  
ttgcctgact ccccgctcgt tagataacta cgatacggga gggcttacca tctggcccca 6840  
gtgctgcaat gataccgcga gaccacgct caccggctcc agatttatca gcaataaacc 6900  
agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 6960  
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 7020  
ttgttgccat tgetgcagc atcgtgggtg cacgctcgtc gtttgggtat gcttcattca 7080  
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 7140  
ttagctcctt cggctcctcg atcgttgtca gaagtaagtt ggcgcagtg ttatcactca 7200



```

tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 7260
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct 7320
cttgcccggc gtcaacacgg gataataccg cgccacatag cagaacttta aaagtgtca 7380
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 7440
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 7500
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 7560
ggaaatggtg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 7620
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacia ataggggttc 7680
cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat 7740
taacctataa aaataggcgt atcacgaggc cctttcgtct tcaagaattg tt 7792

```

&lt;210&gt; 47

&lt;211&gt; 7828

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 47

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttggtgtg cgccggtgta cacaggaagt 240
gacaattttc gcgcggtttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttgggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccagggtg tttttctcag gtgttttccg cgttccgggt caaagttggc gttttattat 480
tatagtcagt acgtaccagt gcaactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcac 960
gctattacca tgggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccaccccat gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggg 1140
aggcgtgtac ggtgggagggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcggccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctggggct ggcgctgggg caccatgctg ctgggcatgc 1380
tgatgatctg cagcgccacc gagaagctgt ggggtgaccgt gtactacggc gtgcccgtgt 1440
ggaaggaggc caccaccacc ctgctctgcg ccagcgacgc caaggcctac gacaccgagg 1500
tgcaaacgt gtgggccacc cagcctgcg tgcccaccga cccaacccc caggaggtgg 1560
tgctggtgaa cgtgaccgag aacttcgaca tgtggaagaa cgacatggtg gagcagatgc 1620
acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcbtgaag ctgaccccc 1680
tgtgcgtgag cctgaagtgc accgacgcta gcaagaactg cagcttcaac atcagacca 1740
gcatccgcg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 1800
tcgacaacga caccaccagc tacagcctga ccagctgcaa caccagcgtg atcaccagg 1860
cctgcccaca ggtgagcttc gagcccatcc ccaaccacta ctgcgcccc gcccgttcg 1920
ccatcctgaa gtgcaaggac aagaagttca acggcaaggg cccctgcacc aacgtgagca 1980

```

ccgtgcagtg caccacacggc atccgccccg tgggtgagcac ccagctgctg ctgaacggta 2040  
 gcctggccga ggaggaggtg gtgatccgca gcgctaactt cgccgacaac gccaaaggta 2100  
 tcatcgtgca gctgaacgag agcgtggaga tcaactgcac ccgcgctagc gccactgca 2160  
 acctgagccg cgccaagtgg aacgacaccc tgaacaagat cgtgatcaag ctgcgcgagc 2220  
 agttcggcaa caagaccatc gtgttcaagc acagcagcgg cggcgacccc gagatcgtga 2280  
 cccacagctt caactgcggc ggcgagttct tctactgcaa cagcgctagc ctgccctgcc 2340  
 gcatcaagca gatcatcaac atgtggcaga aggtgggcaa ggccatgtac gccccccca 2400  
 tcagcggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc cgcgacggcg 2460  
 gcaacagcaa caacgagagc gagatcttcc gtccggggcg cggcgacatg cgcgacaact 2520  
 ggcgcagcga gctgtacaag tacaagggtg tgaagatcga gcccctgggc gtggcccca 2580  
 ccaaggccaa gcttaccgtc caggccccgc agctgctgag cggcatcgtg cagcagcaga 2640  
 acaacctgct gcgcgccatc gaggcccagc agcacctgct gcagctgacc gtgtggggca 2700  
 tcaagcagct gcaggcccgc accctggccg tggagcgcta cctgaaggac cagcagcagc 2760  
 tcgagcagat ctggaaccac accacctgga tggagtggga ccgcgagatc aacaactaca 2820  
 ccagcctgat ccacagcctg atcgaggaga gccagaacca gcaggagaag aacgagcagg 2880  
 agctgctgga gctggacaag tgggcccagc tgtggaactg gttcaacatc accaactggc 2940  
 tgtggtgagg atccagatct gctgtgcctt ctagtgtcca gccatctgtt gtttgccct 3000  
 ccccgctgcc ttccttgacc ctggaagggt ccactcccac tgtcctttcc taataaaatg 3060  
 aggaaattgc atcgcatgtt ctgagtaggt gtcattctat tctggggggt ggggtggggc 3120  
 agcacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct 3180  
 ctatgggtac ccagggccgc ataacttcgt ataattgatg ctatacgaag ttataagatc 3240  
 tgtactgaaa tgtgtgggcg tggcttaagg gtgggaaaga atataaagg tgggggtctt 3300  
 atgtagtctt gtatctgttt tgcagcagcc gccgccgcca tgagcaccaa ctcgtttgat 3360  
 ggaagcattg tgagctcata tttgacaacg cgcattgccc catgggcccg ggtgcgtcag 3420  
 aatgtgatgg gctccagcat tgatggctgc cccgtcctgc ccgcaaactc tactacctg 3480  
 acctacgaga ccgtgtctgg aacgccgttg gagactgcag cctccgcgcg cgttcagcc 3540  
 gctgcagcca ccgcccgcgg gattgtgact gactttgctt tccctgagccc gcttgcaagc 3600  
 agtgcagctt cccgttcac cgcgcgcgat gacaagttga cggctctttt ggcacaattg 3660  
 gattctttga cccgggaact taatgtcgtt tctcagcagc tgttggtatc gcgccagcag 3720  
 gtttctgccc tgaaggcttc ctcccctccc aatgcggttt aaaacataaa taaaaaacca 3780  
 gactctgttt ggatttggtt caagcaagtg tcttgctgtc tttatttagg ggttttgcc 3840  
 gcgcggtagg cccgggacca gcggtctcgg tgcgtgaggg tccctgtgtat tttttccagg 3900  
 acgtggtaaa ggtgactctg gatgttcaga tacatgggca taagcccgtc tctggggtgg 3960  
 aggtagcacc actgcagagc ttcattgctgc ggggtggtgt tgtagatgat ccagtcgtag 4020  
 caggagcgtt gggcgtgggt cctaaaaatg tctttcagta gcaagctgat tgccaggggc 4080  
 aggcccttgg tgaagtgtt taaaaagcgg ttaagctggg atgggtgcat acgtggggat 4140  
 atgagatgca tcttggtact tatttttagg ttggctatgt tcccagccat atccctccgg 4200  
 ggattcatgt tgtgcagaac caccagcaca gtgtatccgg tgcacttggg aaatttgtca 4260  
 tctagcttag aaggaaatgc gtggaagaac ttggagacgc ccttgtgacc tccaagattt 4320  
 tccatgcatt cgtccataat gatggcaatg ggcccacggg cggcgccctg ggcaagata 4380  
 tttctgggat cactaacgtc atagtgtgt tccaggatga gatcgtcata ggccattttt 4440  
 acaaagcgcg ggcggagggt gccagactgc ggtataatgg tccatccgg cccaggggcg 4500  
 tagttaccct cacagatttg catttcccac gctttgagtt cagatggggg gatcatgtct 4560  
 acctgcgggg cgatgaagaa aacggtttcc ggggtagggg agatcagctg ggaagaaagc 4620  
 aggttccctga gcagctgcga cttaccgcag ccggtgggccc cgtaaatacac acctattacc 4680  
 ggctgcaact ggtagttaag agagctgcag ctgccgtcat ccctgagcag gggggccact 4740  
 tcgttaagca tgtccctgac tcgcattgtt tccctgacca aatccgccag aaggcgtcg 4800  
 ccgcccagcg atagcagttc ttgcaaggaa gcaaagtttt tcaacgggtt gagaccgtcc 4860  
 gccgtaggca tgccttttgag cgtttgacca agcagttcca ggcggtccca cagctcggtc 4920  
 acctgctcta cggcatctcg atccagcata tctcctcgtt tcgcgggttg gggcggtttt 4980  
 cgctgtacgg cagtagtcgg tgctcgtcca gacgggccag ggtcatgtct tccacgggc 5040  
 gcagggtcct cgtcagcgta gtctgggtca cgggtgaagg gtgcgtccg ggctgcgcgc 5100  
 tggccagggt gcgcttgagg ctggctcctgc tgggtgctgaa gcgctgccgg tcttcgccct 5160  
 gcgcgtcggc caggtagcat ttgaccatgg tgtcatagtc cagccctcc gcggcgtggc 5220  
 ccttggcgcg cagcttgccc ttggaggagg cgccgcacga ggggcagtg agacttttga 5280

```

gggcgtagag cttggggcgcg agaaataaccg attccgggga gtaggcattcc gcgccgcagg 5340
ccccgcagac ggtctcgcat tccacgagcc aggtgagctc tggccggttcg ggggtcaaaaa 5400
ccagggtttcc cccatgcttt ttgatgcgtt tcttacctct ggtttccatg agccggtgtc 5460
cacgctcggg gacgaaaagg ctgtccgtgt ccccgataac agacttgaga ggcctgtcct 5520
cgagcgggtg tccgcgggtc tctcgtata gaaactcgga ccactctgag acaaaggctc 5580
gcgtccaggc cagcacgaag gaggctaagt gggaggggta gcggtcgttg tccactaggg 5640
ggtccactcg ctccagggtg tgaagacaca tgtcgccctc ttcggcatca aggaagggtg 5700
ttggttttga ggtgtaggcc acgtgaccgg gtgttcctga aggggggcta taaaaggggg 5760
tgggggcgcg ttcgtcctca ctctcttccg catcgctgtc tgcgagggcc agctgttggg 5820
gtgagtcgac gcgaggctgg atggccttcc ccattatgat tcttctcgct tccggcggca 5880
tcgggatgcc cgcgttgag gccatgctgt ccaggcagg agatgacgac catcaggggc 5940
agcttcaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 6000
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 6060
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc 6120
ctgttccgac cgtgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 6180
cgcttttctca tagctcacgc tgtaggatc tcagttcggg gtaggtcgtt cgctccaagc 6240
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 6300
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca 6360
ggattagcag agcgagggtat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact 6420
acggctacac tagaaggaca gtatttggtt tctgcgctct gctgaagcca gttaccttcg 6480
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 6540
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 6600
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggtt ttgggtcatg 6660
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 6720
tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaatc agtgaggcac 6780
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtctgttaga 6840
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6900
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 6960
gaagtgggtc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 7020
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct gcaggcatcg 7080
tggtgtcacg ctctcgtttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 7140
gagttacatg atcccccatg ttgtgcaaaa aagcgggttag ctcttctcgg tctccgatcg 7200
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 7260
ctcttactgt catgccatcc gtaagatgct tttctgtgac tgggtgagtac tcaaccaagt 7320
cattctgaga atagtgtatg cggcgaccga gttgctcttg ccggcgctca acacgggata 7380
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 7440
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacc actcgtgcac 7500
ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 7560
ggcaaaatgc cgcaaaaaag ggaataaggc cgacacggaa atggtgaata ctcatactct 7620
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7680
ttgaatgtat ttagaaaaat aaacaaatag ggggttccgcg cacatttccc cgaaaagtgc 7740
cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7800
cgaggccctt tcgtcttcaa gaattgtt 7828

```

&lt;210&gt; 48

&lt;211&gt; 7789

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 48



ttaattaacc	gcaattctca	tgtttgacag	cttatcatca	tcaataatat	accttatttt	60
ggattgaagc	caatatgata	atgagggggg	ggagtttggt	acgtggcgcg	gggcgtggga	120
acggggcggg	tgacgtagta	gtgtggcgga	agtgtgatgt	tgcaagtgtg	gcggaacaca	180
tgtaagcgac	ggatgtggca	aaagtgcgt	ttttgggtgt	cgccggtgta	cacaggaagt	240
gacaattttc	gcgcgggttt	aggcggatgt	tgtagtaaat	ttgggcgtaa	ccgagtaaga	300
tttggccatt	ttcgcgggaa	aactgaataa	gaggaagtga	aatctgaata	attttgtgtt	360
actcatagcg	cgtaatat	gtctagggcc	gcggggactt	tgaccgttta	cgtggagact	420
cgcccaggtg	tttttctcag	gtgttttccg	cgttccgggt	caaagttagc	gttttattat	480
tatagtcagt	acgtaccagt	gcactggcct	agagcgggcc	cattgcatac	gttgtatcca	540
tatcataata	tgtacattta	tattggctca	tgtccaacat	taccgccatg	ttgacattga	600
ttattgacta	gttattaata	gtaatcaatt	acggggtcct	tagttcatag	cccatatatg	660
gagttccgcg	ttacataact	tacggtaaat	ggcccgcctg	gctgaccgcc	caacgacccc	720
cgcccattga	cgtcaataat	gacgtatgtt	cccatagtaa	cgccaatagg	gactttccat	780
tgacgtcaat	gggtggagta	tttacggtaa	actgccact	tggcagtaca	tcaagtgtat	840
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggccgcg	ctggcattat	900
gcccagtaaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcate	960
gctattacca	tgggtgatgcg	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	1020
tcacggggat	ttccaagtct	ccaccccat	gacgtcaatg	ggagtttggt	ttggcaccaa	1080
aatcaacggg	actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggcggg	1140
aggcgtgtac	gggtgggagg	ctatataagc	agagctcggt	tagtgaaccg	tcagatcgcc	1200
tggagacgcc	atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	1260
cgtcacgcgt	gtcgacacgt	gtgatcagat	atcgcgcccg	ctctagaatg	cgcgtaagg	1320
agaagtacca	gcacctgtgg	cgctggggct	ggcgctgggg	caccatgctg	ctgggcatgc	1380
tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgcacaacgt	gtgggccacc	cacgcctgcg	tgcccaccga	ccccaacccc	caggagggtg	1560
tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgctggaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacctga	agaacgacac	caacaccaac	agcagcagcg	1740
gccgcatgat	catggagaag	ggcgagatca	agaactgcag	cttcgctagc	accagctgca	1800
acaccagcgt	gatcaccag	gcctgcccc	agggtgagct	cgagcccatc	cccaaccact	1860
actgcgcccc	cgccggcttc	gccatcctga	agtgcaggga	caagaagttc	aacggcaagg	1920
gcccctgcac	caacgtgagc	accgtgcagt	gcacccacgg	catccgcccc	gtggtgagca	1980
cccagctgct	gctgaacggg	agcctggccg	aggaggagg	ggtgatccgc	agcgctaact	2040
tgcgagacaa	cgccaagggt	atcatcgtgc	agctgaacga	gagcgtggag	atcaactgca	2100
cccgcgctag	cgccactgc	aacctgagcc	gcgccaagtg	gaacgacacc	ctgaacaaga	2160
tcgatgatcaa	gctgcgcgag	cagttcggca	acaagaccat	cgtgttcaag	cacagcagcg	2220
gcggcgaccc	cgagatcgtg	acccacagct	tcaactgcgg	cggcgagttc	ttctactgca	2280
acagcgctag	cctgccctgc	cgcataagc	agatcatcaa	catgtggcag	aagggtggga	2340
aggccatgta	cgcccccccc	atcagcgggc	agatccgctg	cagcagcaac	atcaccggcc	2400
tgctgctgac	ccgcgacggc	ggcaacagca	acaacgagag	cgagatcttc	cgtccggggc	2460
gcggcgacat	gcgcgacaa	tggcgacagc	agctgtacaa	gtacaagggt	gtgaagatcg	2520
agcccctggg	cgtggccccc	accaaggcca	agcttaccgt	ccaggcccg	cagctgctga	2580
gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggcccg	cagcacctgc	2640
tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccg	cacctggcc	gtggagcgt	2700
acctgaagga	ccagcagcag	ctcgagcaga	tctggaacca	caccacctgg	atggagtggg	2760
accgcgagat	caacaactac	accagcctga	tccacagcct	gatcgaggag	agccagaacc	2820
agcaggagaa	gaacgagcag	gagctgctgg	agctggacaa	gtgggcccagc	ctgtggaact	2880
ggttcaacat	caccaactgg	ctgtggtgag	gatccagatc	tgtgtgcct	tctagtgtcc	2940
agccatctgt	tgtttgcccc	tccccgtgc	cttccctgac	cctggaagg	gccactcca	3000
ctgtcccttc	ctaataaaat	gaggaaattg	catcgcat	tctgagtagg	tgtcattcta	3060
ttctgggggg	tgggggtggg	cagcacagca	agggggagga	ttgggaagac	aatagcaggc	3120
atgctgggga	tgcgggtggg	tctatgggta	cccaggggcc	cataacttcg	tataatgtat	3180
gctatacgaa	gttataagat	ctgtactgaa	atgtgtgggc	gtggcttaag	ggtgggaaag	3240
aatatataag	gtgggggtct	tatgtagttt	tgtatctgtt	ttgcagcagc	cgccgcccgc	3300

atgagcacca actcgtttga tggaaagcatt gtgagctcat atttgacaac gcgcatgccc 3360  
 ccatggggccg ggggtgcgtca gaatgtgatg ggctccagca ttgatggtcg ccccgctcctg 3420  
 cccgcaaact ctactacctt gacctacgag accgtgtctg gaacgccgtt ggagactgca 3480  
 gcctccgccc ccgcttcagc cgctgcagcc accgcccgcg ggattgtgac tgactttgct 3540  
 ttcttgagcc cgcttgcaag cagtgcagct tcccgttcat ccgcccgcga tgacaagttg 3600  
 acggctcttt tggcacaatt ggattctttg acccggaac ttaatgtcgt ttctcagcag 3660  
 ctgttggatc tgcgccagca ggtttctgcc ctgaaggctt cctccccctc caatgcgggt 3720  
 taaaacataa ataaaaaacc agactctgtt tggatttggg tcaagcaagt gtcttgctgt 3780  
 ctttatttag gggttttgcg cgcgcggtag gcccgggacc agcggtctcg gtcgttgagg 3840  
 gtctgtgta ttttttccag gacgtggtaa aggtgactct ggatgttcag atacatgggc 3900  
 ataagcccgt ctctgggggtg gaggtagcac cactgcagag cttcatgctg cgggggtggg 3960  
 ttgtagatga tccagtcgta gcaggagcgc tgggcgtggg gcctaaaaat gtctttcagt 4020  
 agcaagctga ttgccagggg caggcccttg gtgtaagtgt ttacaaagcg gtttaagctg 4080  
 gatgggtgca tacgtgggga tatgagatgc atcttggact gtatttttag gttggctatg 4140  
 ttcccagcca tatccctccg gggattcatg ttgtgcagaa ccaccagcac agtgtatccg 4200  
 gtgcacttgg gaaatttgtc atgtagctta gaaggaaatg cgtggaagaa cttggagacg 4260  
 cccttgtgac ctccaagatt ttccatgcat tcgtccataa tgatggcaat gggcccacgg 4320  
 gcggcgccct gggcgaagat atttctggga tctaactcgt catagtgtg ttccaggatg 4380  
 agatcgctcat aggcattttt taaaaagcgc gggcgagggg tgccagactg cgggtataatg 4440  
 gttccatccg gccagggggc gtagttaccc tcacagattt gcatttccca cgctttgagt 4500  
 tcagatgggg ggatcatgtc tacctgcggg gcgatgaaga aaacgggtttc cggggtaggg 4560  
 gagatcagct gggaagaaag caggttcctg agcagctgcg acttaccgca gccgggtggg 4620  
 ccgtaaatca cacctattac cggctgcaac tggtagttaa gagagctgca gctgccgtca 4680  
 tccctgagca ggggggccac ttctgtaagc atgtccctga ctgcgatgtt ttccctgacc 4740  
 aaatccgcca gaaggcgtc gccgcccagc gatagcagtt cttgcaagga agcaaagttt 4800  
 ttcaacgggt tgagaccgtc cgccgtaggc atgcttttga gcgtttgacc aagcagttcc 4860  
 aggcgggtccc acagctcggg cacctgctct acggcatctc gatccagcat atctcctcgt 4920  
 ttctcggggtt ggggcggctt tcgctgtacg gcagtagtcg gtgctcgtcc agacggggca 4980  
 gggtcatgtc ttccacaggg cgcagggtcc tcgtcagcgt agtctgggtc acggtgaagg 5040  
 ggtgcgctcc gggctgcgcg ctggccaggg tgcgcttgag gctggtcctg ctggtgctga 5100  
 agcgtgcgcg gtcttcgccc tgcgcgtcgg ccaggtagca tttgaccatg gtgtcatagt 5160  
 ccagcccctc cgcggcgtgg cccttgggcg gcagcttgcc cttggaggag gcgcccacg 5220  
 aggggcagtg cagacttttg agggcgtaga gcttgggcgc gagaaatacc gattccgggg 5280  
 agtaggcata cgcgcgcag gcccgcgaga cggctctcga ttccacgagc caggtagact 5340  
 ctggccgttc ggggtcaaaa accagggttc ccccatgctt tttgatgctt ttcttacctc 5400  
 tggtttccat gaggcgggtg ccacgctcgg tgacgaaaag gctgtccgtg tccccgtata 5460  
 cagacttgag aggcctgtcc tcgagcgggt ttccgcgggt ctcctcgtat agaaactcgg 5520  
 accactctga gacaaaggct cgcgtccagg ccagcacgaa ggaggctaag tgggaggggt 5580  
 agcggtcgtt gtccactagg gggteccact gctccagggt gtgaagacac atgtcgccct 5640  
 ctctcgcatc aaggaagggt attggtttgt aggtgtaggc cacgtgaccg ggtgttcctg 5700  
 aaggggggct ataaaagggt gtggggggcg gttcgtcctc actctcttcc gcacgctgt 5760  
 ctgcgagggc cagctgttgg ggtgagtcga cgcgaggctg gatggccttc cccattatga 5820  
 ttcttctcgc ttccggcggc atcgggatgc ccgcttgca ggccatgctg tccaggcagg 5880  
 tagatgacga ccatcaggga cagcttcaag gccagcaaaa ggccagggaac cgtaaaaagg 5940  
 ccgcttctgt ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 6000  
 gctcaagtca gaggtggcga aaccgcagag gactataaag ataccaggcg tttccccctg 6060  
 gaagctccct cgtgcgctct cctgttccga cctgcccgt taccggatac ctgtccgcct 6120  
 ttctcccttc ggggaagcgt gcgctttctc atagctcacg ctgtagggtat ctgagttcgg 6180  
 tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc cccggttcag cccgaccgct 6240  
 gcgccttata cggttaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 6300  
 tggcagcagc cactggtaac aggttagca gagcgaggta tgtaggcggg gctacagagt 6360  
 tcttgaagt gtggcctaac tacggctaca ctagaaggac agtatttggg atctgcgctc 6420  
 tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 6480  
 ccgctggtag cgggtggttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat 6540  
 ctcaagaaga tcctttgatc tttctacgg ggtctgacgc tcagtggaaac gaaaactcac 6600

```

gttaagggat tttgggtcatg agattatcaa aaaggatcctt cacctagatc ctttttaaatt 6660
aaaaatgaag ttttaaataca atctaaagta tatatgagta aacttgggtct gacagttacc 6720
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 6780
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 6840
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 6900
cagccggaag ggccgagcgc agaagtgggc ctgcaacttt atccgcctcc atccagtcta 6960
ttaattggtg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacggtg 7020
ttgccattgc tgcaggcatc gtgggtgtcac gctcgtcgtt tggatatggct tcattcagct 7080
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 7140
gctccttcgg tcttcgatac gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg 7200
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 7260
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 7320
gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 7380
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 7440
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 7500
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggataaagg gcgacacgga 7560
aatgttgaat actcatactc ttcttttttc aatattattg aagcatttat cagggttatt 7620
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 7680
gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa 7740
cctataaaaa taggcgtatc acgaggccct ttcgtcttca agaattggtt 7789

```

&lt;210&gt; 49

&lt;211&gt; 7711

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pAdApt containing HIV genes

&lt;400&gt; 49

```

ttaattaacc gcaattctca tgtttgacag cttatcatca tcaataatat accttatttt 60
ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcgtggga 120
acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgtg cgccgggtgta cacaggaagt 240
gacaattttc gcgcgggttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
cgcccagggt tttttctcag gtgttttcgc cgttccgggt caaagttagc gttttattat 480
tatagtcagt acgtaccagt gcactggcct agagcggccc cattgcatac gttgtatcca 540
tatcataata tgtacattta tattggetca tgtccaacat taccgccatg ttgacattga 600
ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag cccatatatg 660
gagttccgcg ttacataact tacggtaa at ggcccgcctg gctgaccgcc caacgacccc 720
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 780
tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 840
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 900
gccagttaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 960
gctattacca tgggtgatgc gttttggcag tacatcaatg ggcgtggata gcggtttgac 1020
tcacggggat ttccaagtct ccaccccatg gacgtcaatg ggagtttgtt ttggcaccaa 1080
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggg 1140
aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc 1200
tggagacgcc atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc 1260
cgtcaccgtc gtcgacacgt gtgatcagat atcgcgcccg ctctagaatg cgcgtgaagg 1320
agaagtacca gcacctgtgg cgctgggggt ggcgctgggg caccatgctg ctgggcatgc 1380

```



tgatgatctg	cagcgccacc	gagaagctgt	gggtgaccgt	gtactacggc	gtgcccgtgt	1440
ggaaggaggc	caccaccacc	ctgctctgcg	ccagcgacgc	caaggcctac	gacaccgagg	1500
tgcacaacgt	gtggggccacc	cacgcctgcg	tgcccaccga	ccccaacccc	caggagggtg	1560
tgctggtgaa	cgtgaccgag	aacttcgaca	tgtggaagaa	cgacatggtg	gagcagatgc	1620
acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	ctgaccccc	1680
tgtgcgtgag	cctgaagtgc	accgacgcta	gcaccagctg	caacaccagc	gtgatcacc	1740
aggcctgccc	caagggtgagc	ttcgagccca	tccccaacca	ctactgcgcc	cccgcgggt	1800
tcgccatcct	gaagtgcaag	gacaagaagt	tcaacggcaa	gggcccctgc	accaacgtga	1860
gcaccgtgca	gtgcaccac	ggcatccgcc	ccgtggtgag	caccagctg	ctgctgaacg	1920
gtagcctggc	cgaggaggag	gtggtgatcc	gcagcgctaa	cttcgccgac	aacgccaagg	1980
tgatcatcgt	gcagctgaac	gagagcgtgg	agatcaactg	caccgcgct	agcgcctact	2040
gcaacctgag	ccgcgccaa	tggaaacgaca	ccctgaacaa	gatcgtgatc	aagctgcgcg	2100
agcagttcgg	caacaagacc	atcgtgttca	agcacagcag	cgccggcgac	cccgagatcg	2160
tgaccacag	cttcaactgc	ggcgccgagt	tcttctactg	caacagcgct	agcctgccct	2220
gccgcatcaa	gcagatcatc	aacatgtggc	agaagggtgg	caaggccatg	tacgcccc	2280
ccatcagcgg	ccagatccgc	tgcagcagca	acatcaccgg	cctgctgctg	accgcgcacg	2340
gcggcaacag	caacaacgag	agcgagatct	tccgtccggg	cgccggcgac	atgcgcgaca	2400
actggcgag	cgagctgtac	aagtacaagg	tgggtgaagat	cgagcccctg	ggcgtggccc	2460
ccaccaaggc	caagcttacc	gtccaggccc	gccagctgct	gagcggcatc	gtgcagcagc	2520
agaacaacct	gctgcgcgcc	atcgaggccc	agcagcacct	gctgcagctg	accgtgtggg	2580
gcatcaagca	gctgcaggcc	cgcaccctgg	ccgtggagcg	ctacctgaag	gaccagcagc	2640
agctcgagca	gatctggaac	cacaccacct	ggatggagtg	ggaccgcgag	atcaacaact	2700
acaccagcct	gatccacagc	ctgatcgagg	agagccagaa	ccagcaggag	aagaacgagc	2760
aggagctgct	ggagctggac	aagtgggcca	gcctgtggaa	ctggttcaac	atcaccaact	2820
ggctgtggtg	aggatccaga	tctgctgtgc	cttctagtgt	ccagccatct	gttgtttgcc	2880
cctccccgt	gccttccttg	accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	2940
atgaggaaat	tgcategcat	tgtctgagta	ggtgtcattc	tattctgggg	ggtggggtgg	3000
ggcagcacag	caagggggag	gattgggaag	acaatagcag	gcattgctggg	gatgcggtgg	3060
gctctatggg	taccaggggc	cgcataactt	cgtataatgt	atgctatacg	aagttataag	3120
atctgtactg	aaatgtgtgg	gcgtggctta	aggggtggaa	agaatatata	aggtgggggt	3180
cttatgtagt	tttgtatctg	ttttgcagca	gccgccgcgc	ccatgagcac	caactcgttt	3240
gatggaagca	ttgtgagctc	atatttgaca	acgcgcagtc	cccatgggc	cgggggtgct	3300
cagaatgtga	tgggctccag	cattgatggg	cgccccgtcc	tggccgcaaa	ctctactacc	3360
ttgacctacg	agaccgtgtc	tggaaacgcg	ttggagactg	cagcctccgc	cgcgccttca	3420
gccgctgcag	ccaccgcccg	cgggattgtg	actgactttg	ctttcctgag	cccgccttgc	3480
agcagtgcag	cttcccgttc	atccgcccgc	gatgacaagt	tgacggctct	tttggcacia	3540
ttggattctt	tgaccgggga	acttaatgtc	gtttctcagc	agctgttgga	tctgcgccag	3600
caggtttctg	ccctgaaggc	ttcctccct	cccaatgcgg	tttaaaacat	aaataaaaaa	3660
ccagactctg	tttggatttg	gatcaagcaa	gtgtcttgct	gtctttattt	aggggttttg	3720
cgcgcgcggt	aggcccggga	ccagcggctc	cggctcgttg	gggtcctgtg	tattttttcc	3780
aggacgtggg	aaagggtgact	ctggatgttc	agatacatgg	gcataagccc	gtctctgggg	3840
tggaggtagc	accactgcag	agcttcatgc	tgcgggggtg	tgttgtagat	gatccagtcg	3900
tagcaggagc	gctgggctg	gtgcctaaaa	atgtctttca	gtagcaagct	gattgccagg	3960
ggcaggccct	tgggtgtaagt	gtttacaaag	cgggttaagct	gggatgggtg	catacgtggg	4020
gatatgagat	gcattcttga	ctgtattttt	aggttggtta	tgttcccagc	catatccctc	4080
cggggattca	tgttggtgcag	aaccaccagc	acagtgtatc	cgggtgcactt	gggaaatttg	4140
tcatgtagct	tagaaggaaa	tgcgtggaag	aacttgagga	cgcccttggtg	acctccaaga	4200
ttttccatgc	attcgtccat	aatgatggca	atgggcccac	gggcggcggc	ctgggcgaag	4260
atatttctgg	gatcactaac	gtcatagttg	tgttccagga	tgagatcgtc	ataggccatt	4320
tttaciaaagc	gcgggcggag	ggtgccagac	tgcggtataa	tggttccatc	cggcccaggg	4380
gcgtagttac	cctcacagat	ttgcatttcc	cacgctttga	gttcagatgg	ggggatcatg	4440
tctacctgcg	gggcgatgaa	gaaaacggtt	tccggggtag	gggagatcag	ctgggaagaa	4500
agcaggttcc	tgagcagctg	cgacttaccg	cagccggtgg	gcccgtaa	cacacctatt	4560
accggctgca	actggtagtt	aagagagctg	cagctgccgt	catccctgag	cagggggggc	4620
acttcgttaa	gcattgtccct	gactcgcag	ttttccctga	ccaaatccgc	cagaaggcgc	4680

tcgccgcca	gcgatagcag	ttcttgcaag	gaagcaaagt	ttttcaacgg	tttgagaccg	4740
tccgccgtag	gcatgctttt	gagcgtttga	ccaagcagtt	ccaggcggtc	ccacagctcg	4800
gtcacctgct	ctacggcatc	tcgatccagc	atatctcttc	gtttcgcggg	ttggggcggc	4860
tttcgctgta	cggcagtagt	cggtgctcgt	ccagacgggc	cagggtcatg	tctttccacg	4920
ggcgcagggt	cctcgtcagc	gtagtctggg	tcacggtgaa	ggggtgcgct	ccgggctgcg	4980
cgctggccag	ggtgcgcttg	aggctgggtc	tgctgggtgct	gaagcgctgc	cgggtcttcgc	5040
cctgcgcgtc	ggccaggtag	catttgacca	tggtgtcata	gtccagcccc	tccgcggcgt	5100
ggcccttgge	gcgcagcttg	cccttgagg	aggcgccgca	cagggggcag	tgcagacttt	5160
tgagggcgta	gagcttgggc	gcgagaaata	ccgattccgg	ggagtaggca	tccgcgccgc	5220
aggccccgca	gacgggtctcg	cattccacga	gccagggtgag	ctctggccgt	tcgggggtcaa	5280
aaaccagggt	tcccccatgc	tttttgatgc	gtttcttacc	tctgggtttcc	atgagccgggt	5340
gtccacgctc	ggtgacgaaa	aggctgtccg	tgtccccgta	tacagacttg	agaggcctgt	5400
cctcgagcgg	tggtccgcgg	tcctcctcgt	atagaaactc	ggaccactct	gagacaaagg	5460
ctcgcgtcca	ggccagcacg	aaggaggcta	agtgggaggg	gtagcggtcg	ttgtccacta	5520
gggggtccac	tcgctccagg	gtgtgaagac	acatgtcgcc	ctcttcggca	tcaagggaagg	5580
tgattgggtt	gtagggtgtag	gccacgtgac	cgggtgttcc	tgaagggggg	ctataaaagg	5640
gggtgggggc	gcgttcgtcc	tcactctctt	ccgcctcgtc	gtctgcgagg	gccagctgtt	5700
gggggtgagtc	gacgcgagge	tggatggcct	tccccattat	gattcttctc	gcttccggcg	5760
gcctcgggat	gccccgcttg	caggccatgc	tgtccaggca	ggtagatgac	gaccatcagg	5820
gacagcttca	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcggttt	5880
tccataggct	ccgccccctt	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	5940
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaaagctcc	ctcgtgcgct	6000
ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	6060
tggecgcttc	tcatagctca	cgtgttaggt	atctcagttc	ggtgtagggtc	gttcgctcca	6120
agctgggctg	tgtgcacgaa	cccccggttc	agcccgaccg	ctgcgcctta	tccggtaact	6180
atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	6240
acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	6300
actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	6360
tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaac	caccgctgggt	agcggtggtt	6420
tttttggttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gaccccttga	6480
tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	acgttaagggt	atthttgggtca	6540
tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	agtttttaa	6600
caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	atcagtgagg	6660
cacctatctc	agcgatctgt	ctatttcgtt	catccatagt	tgcttgactc	cccgtcgtgt	6720
agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	ataccgcgag	6780
acccacgctc	accggctcca	gatttatcag	caataaacca	gccagccgga	agggccgagc	6840
gcagaagtgg	tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	tgccgggaag	6900
ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	tggttgccatt	gctgcaggca	6960
tcgtgggtgc	acgctcgtcg	tttgggtatg	cttcattcag	ctccgggttc	caacgatcaa	7020
ggcgagttac	atgatcccc	atgttggtgca	aaaaagcggt	tagctccttc	ggctcctcga	7080
tcgttggtcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	gactgcata	7140
attctcttac	tgtcatgcca	tccgtaagat	gcttttctgt	gactgggtgag	tactcaacca	7200
agtcattctg	agaatagtg	atgcggcgac	cgagttgctc	ttgcccggcg	tcaacacggg	7260
ataataccgc	gccacatagc	agaactttta	aagtgtcat	cattggaaaa	cgttcttcgg	7320
ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	cccactcgtg	7380
cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	gcaaaaacag	7440
gaaggcaaaa	tgccgcaaaa	aagggaataa	gggcgacacg	gaaatggtga	atactcatac	7500
tcttcctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	agcggataca	7560
tatttgaaatg	tatttagaaa	aataaaca	taggggttcc	gcgcacattt	ccccgaaaag	7620
tgccacctga	cgtctaagaa	accattatta	tcatgacatt	aacctataaa	aataggcgta	7680
tcacgaggcc	ctttcgtctt	caagaattgt	t			7711

&lt;211&gt; 6577

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 50

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgaac taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatatata acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcgggtg tgttaacggg ggaggcgagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcaagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag gctagcagcc tgaagtgcac cgacctgaag 2040
aacgacacca acaccaacag cagcagcggc cgcagatca tggagaaggg cgagatcaag 2100
aactgcagct tcaacatcag caccagcatc cgcggcaagg tgcagaagga gtacgccttc 2160
ttctacaagc tggacatcat ccccatcgac aacgacacca ccagctacag cctgaccagc 2220
tgcaacacca gcgtgatcac ccaggcctgc cccaagggtg gcttcgagcc catccccaac 2280
cactactgcg ccccgccgg ctctcgccatc ctgaagtgc aaggacaaga gttcaacggc 2340
aaggggccct gcaccaacgt gagcaccgtg cagtgcaccc acggcatccg ccccggtggtg 2400
agcaccagc tgctgctgaa cggtagcctg gccgaggagg aggtgggtgat ccgcagcgt 2460
aacttcgccc acaacgccaa ggtgatcatc gtgcagctga acgagagcgt ggagatcaac 2520
tgaccccgcc ccaacaacaa caccgcgaag agcatccaca tcggccccgg ccgcgccttc 2580
tacaccaccg gcgagatcat cggcgacatc cgccaggccc actgcaacct gagccgcgcc 2640
aagtggaacg acaccctgaa caagatcgtg atcaagctgc gcgagcagtt cggcaacaag 2700
accatcgtgt tcaagcacag cagcggcggc gaccccgaga tcgtgacca cagcttcaac 2760
tgcggcggcg agttcttcta ctgcaacagc acccagctgt tcaacagcac ctggttcaac 2820

```



agcacctgga gcaccgaggg cagcaacaac accgagggca gcgacaccat caccctgccc 2880  
tgccgcatca agcagatcat caacatgtgg cagaaggtgg gcaaggccat gtacgcccc 2940  
cccatcagcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gaccgcgcgac 3000  
ggcggcaaca gcaacaacga gagcgagatc ttccgtccgg gcggcggcga catgcgcgac 3060  
aactggcgca gcgagctgta caagtacaag gtggtgaaga tcgagcccct gggcgtggcc 3120  
cccaccaagg ccaagcttac cgtccaggcc cgccagctgc tgagcggcat cgtgcagcag 3180  
cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 3240  
ggcatcaagc agctgcaggc ccgcaccctg gccgtggagc gctacctgaa ggaccagcag 3300  
cagctcgagc agatctggaa ccacaccacc tggatggagt gggaccgcga gatcaacaac 3360  
tacaccagcc tgatccacag cctgatcgag gagagccaga accagcacga gaagaacgag 3420  
caggagctgc tggagctgga caagtgggccc agcctgtgga actggttcaa catcaccaac 3480  
tggtctgtgt acatcaagct gttcatcatg atcgtgggccc gcctgggtggg cctgcgcac 3540  
gtgttcgccc tgctgagcat ctgaggatcc agatctgctg tgccttctag ttgccagcca 3600  
tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc 3660  
ctttcctaataaaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg 3720  
gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct 3780  
ggggatgcgg tgggctctat gggtagccag gtgctgaaga attgaccgga ttccctcctg 3840  
gccagaaaga agcaggcaca tccccctctc tgtgacacac cctgtccacg cccctgggtc 3900  
ttagttccag cccactcat aggacactca tagctcagga gggctccgcc ttcaatccca 3960  
cccgtaaaag tacttgagc ggtctctccc tccctcatca gccacccaaa ccaaacctag 4020  
cctccaagag tgggaagaaa ttaagcaag ataggctatt aagtgcagag ggagagaaaa 4080  
tgccccaac atgtgaggaa gtaatgagag aaatcataga attttaaggc catcatggcc 4140  
ttaatcttcc gcttccctgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4200  
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4260  
aaagaacatg tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4320  
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4380  
gaggtggcga aaccgcagag gactataaag ataccaggcg ttccccctg gaagctccct 4440  
cgtgcgctct cctgttccga cctgcccgt taccggatac ctgtccgct ttctccctc 4500  
gggaagcgtg gcgcttctc atagctcacg ctgtaggtat ctgagttcgg tgtaggtcgt 4560  
tcgctccaag ctgggctgtg tgcacgaacc cccggttcag ccgaccgct gcgccttate 4620  
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4680  
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagt 4740  
gtggcctaac tacggctaca ctagaagaac agtatttgggt atctgcgctc tgcgaagcc 4800  
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4860  
cgggtggttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga 4920  
tcctttgatc ttttctacgg ggtctgacgc tcagtgaac gaaaactcac gttaagggat 4980  
tttggctcatg agattatcaa aaaggatctt cacctagatc cttttaaat aaaaatgaag 5040  
ttttaaatca atctaaagta tatatgagta aacttggctt gacagttacc aatgcttaat 5100  
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5160  
gggggggggg cgctgaggct tgcctcgtga agaagggtgt gctgactcat accaggcctg 5220  
aatcgcccca tcatccagcc agaaagttag ggagccacgg ttgatgagag ctttgttgta 5280  
ggtggaccag ttggtgattt tgaacttttg ctttgcacg gaacggctctg cgttgtcggg 5340  
aagatgcgtg atctgatcct tcaactcagc aaaagttcga ttattcaac aaagccggcg 5400  
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaacca ttctgattag 5460  
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5520  
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5580  
atggcaagat cctggatctg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5640  
aatttccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5700  
tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760  
ttacgctcgt catcaaaatc actcgcacat accaaaccgt tattcattcg tgattgcgcc 5820  
tgagcgagac gaaatacgcg atcgtgtta aaaggacaat tacaacagg aatcgaatgc 5880  
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 5940  
tctaatacct ggaatgctgt tttccgggg atcgcagtgg tgagtaacca tgcacatca 6000  
ggagtacgga taaaatgctt gatggctcga agaggcataa attccgtcag ccagtttagt 6060  
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120

```

tctggcgcacat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6180
tcgcgagccc atttataccc atataaatca gcatccatgt tgggaatttaa tcgcggcctc 6240
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct tcccccccc cccattatt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgtat cacgaggccc tttcgtc 6577

```

&lt;210&gt; 51

&lt;211&gt; 6577

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 51

```

tcgcgcggttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggetat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgcctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagb gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgcaa ggcctacgac 2100

```

accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160  
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220  
cagatgcacg aggacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 2280  
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340  
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400  
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460  
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccgc tagcatcaac 2520  
tgcacccgcc ccaacaacaa caccgcgaag agcatccaca tcggccccgg ccgcgccttc 2580  
tacaccaccg gcgagatcat cggcgacatc cgccaggccc actgcaacct gagccgcgcc 2640  
aagtggaaac acaccctgaa caagatcgtg atcaagctgc gcgagcagtt cggcaacaag 2700  
accatcgtgt tcaagcacag cagcggcggc gaccccgaga tcgtgacca cagcttcaac 2760  
tgccggcggc agttcttcta ctgcaacagc acccagctgt tcaacagcac ctggttcaac 2820  
agcacctgga gcaccgaggc cagcaacaac accgagggca gcgacaccat caccctgccc 2880  
tgccgcatac agcagatcat caacatgttg cagaagggtg gcaaggccat gtacgcccc 2940  
cccatcagcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 3000  
ggcggcaaca gcaacaacga gagcgagatc ttccgtccgg gcggcggcga catgcgcgac 3060  
aactggcgca gcgagctgta caagtacaag gtggtgaaga tcgagcccct gggcgtggcc 3120  
cccaccaagg ccaagcttac cgtccaggcc cgccagctgc tgagcggcat cgtgcagcag 3180  
cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 3240  
ggcatcaagc agctgcaggc ccgcaccctg gccgtggagc gctacctgaa ggaccagcag 3300  
cagctcgagc agatctggaa ccacaccacc tggatggagt gggaccgcca gatcaacaac 3360  
tacaccagcc tgatccacag cctgatcgag gagagccaga accagcacga gaagaacgag 3420  
caggagctgc tggagctgga caagtgggccc agcctgtgga actggttcaa catcaccaac 3480  
tggctgtggt acatcaagct gttcatcatg atcgtgggcg gcctggtggg cctgcgcatac 3540  
gtgttcgccg tgctgagcat ctgaggatcc agatctgctg tgccttctag ttgccagcca 3600  
tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aagggtgccac tcccactgtc 3660  
ctttcctaataaaaatgagga aattgcacgc cattgtctga gtaggtgtca ttctattctg 3720  
gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct 3780  
ggggatgcgg tgggctctat gggtagccag gtgctgaaga attgaccggg ttccctcctgg 3840  
gccagaaaga agcaggcaca tccccttctc tgtgacacac cctgtccacg cccctggttc 3900  
ttagttccag cccactcat aggacactca tagctcagga gggctccgcc ttcaatccca 3960  
cccgtataag tacttgagc ggtctctccc tccctcatca gccacccaaa ccaaacctag 4020  
cctccaagag tgggaagaaa tttaaagcaag ataggctatt aagtgcagag ggagagaaaa 4080  
tgctccaac atgtgaggaa gtaatgagag aaatcataga attttaaggc catcatggcc 4140  
ttaatcttcc gcttctctgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4200  
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4260  
aaagaacatg tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4320  
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4380  
gaggtggcga aacccgacag gactataaag ataccaggcg ttccccctg gaagctccct 4440  
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgctt ttctcccttc 4500  
gggaagcgtg gcgcttctc atagctcacg ctgtaggat ctcagttcgg tgtaggctgt 4560  
tcgctccaag ctgggctgtg tgcacgaacc cccggttcag cccgaccgct gcgccttate 4620  
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4680  
cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg 4740  
gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 4800  
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4860  
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4920  
tcctttgatc ttttctacgg ggtctgacgc tcagtggaaac gaaaactcac gttaagggat 4980  
tttggctatg agattatcaa aaaggatctt cacctagatc cttttaaat aaaaatgaag 5040  
ttttaaatca atctaaagta tatatgagta aacttggctc gacagttacc aatgcttaat 5100  
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5160  
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5220  
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5280  
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacgggtctg cgttgcggg 5340  
aagatgcgtg atctgatect tcaactcagc aaaagttcga ttatttcaac aaagccgccg 5400



```

tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5520
tattttttgaa aaagccggtt ctgtaatgaa ggagaaaact caccgaggca gtcccatagg 5580
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5640
aatttcccct cgtcaaaaat aagggttatca agtgagaaat caccatgagt gacgactgaa 5700
tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760
ttacgctcgt catcaaaaat actcgcacat accaaaccgt tattcattcg tgattgcgcc 5820
tgagcgagac gaaatacgcg atcgcgtgtta aaaggacaat tacaaacagg aatcgaatgc 5880
aaccggcgca ggaacactgc cagcgcacat acaatatttt cacctgaatc aggatattct 5940
tctaatacct ggaatgctgt tttcccgggg atcgcagtggt tgagtaacca tgcacatca 6000
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6060
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6180
tcgcgagccc atttataccc atataaatca gcattccatgt tgggaattta tgcgggctc 6240
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct tcccccccc cccattattt gaagcattta tcagggttat 6420
tgtctcatga gcgatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgat caccaggccc tttcgtc 6577

```

&lt;210&gt; 52

&lt;211&gt; 6718

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 52

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380

```

cgcagttttt	attaaacata	gcgtgggato	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgccaccgag	aagctgtggg	tgaccgtgta	ctacggcgtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcca	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtggtgc	tgggtgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacacca	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaagg	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtgag	cttcgagccc	atccccaa	actactgcgc	ccccgccggc	2580
ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	agggcccttg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	acttcgcgga	caacgccaag	2760
gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	gcaccgccc	caacaacaac	2820
acccgcaaga	gcatccacat	cggccccggc	cgcgccttct	acaccaccgg	cgagatcatc	2880
ggcgacatcc	gccaggccca	ctgcaacctg	gctagcttct	actgcaacag	caccagctg	2940
ttcaacagca	cctgggtcaa	cagcacctgg	agcaccgagg	gcagcaacaa	caccgagggc	3000
agcgacacca	tcaccctgcc	ctgcccgcac	aagcagatca	tcaacatgtg	gcagaagggtg	3060
ggcaaggcca	tgtacgcccc	ccccatcagc	ggccagatcc	gctgcagcag	caacatcacc	3120
ggcctgctgc	tgaccgcgga	cggcggcaac	agcaacaacg	agagcgagat	cttccgtccg	3180
ggcggcgggc	acatgcgcga	caactggcgc	agcgagctgt	acaagtacaa	ggtggtgaag	3240
atcgagcccc	tgggcgtggc	ccccaccaag	gccaagctta	ccgtccaggc	ccgccagctg	3300
ctgagcggca	tcgtgcagca	gcagaacaac	ctgctgcgcg	ccatcgaggc	ccagcagcac	3360
ctgctgcagc	tgaccgtgtg	gggcatcaag	cagctgcagg	cccgcaccct	ggccgtggag	3420
cgtacactga	aggaccagca	gcagctcgag	cagatctgga	accacaccac	ctggatggag	3480
tgggaccgcg	agatcaacaa	ctacaccagc	ctgatccaca	gcctgatcga	ggagagccag	3540
aaccagcacg	agaagaacga	gcaggagctg	ctggagctgg	acaagtgggc	cagcctgtgg	3600
aactgggtca	acatcaccaa	ctggctgtgg	tacatcaagc	tgttcatcat	gatcgtgggc	3660
ggcctggtgg	gcctgcgcac	cgtgttcgcc	gtgctgagca	tctgaggatc	cagatctgct	3720
gtgccttcta	gttgccagcc	atctgttgtt	tgccccctcc	ccgtgccttc	cttgaccctg	3780
gaaggtgcca	ctcccactgt	cctttcctaa	taaaatgagg	aaattgcac	gcattgtctg	3840
agtaggtgtc	attctattct	gggggggtggg	gtggggcagg	acagcaaggg	ggaggattgg	3900
gaagacaata	gcaggcatgc	tggggatgcg	gtgggctcta	tgggtacca	ggtgctgaag	3960
aattgacccg	gttcctcctg	ggccagaaa	aagcaggcac	atccccctt	ctgtgacaca	4020
ccctgtccac	gcccctgggt	cttagttcca	gccccactca	taggacactc	atagctcagg	4080
agggctccgc	cttcaatccc	acccgctaaa	gtacttgagg	cggctctctc	ctccctcatc	4140
agcccaccaa	accaaacccta	gcctccaaga	gtgggaagaa	attaaagcaa	gataggctat	4200
taagtgcaga	gggagagaaa	atgcctccaa	catgtgagga	agtaatgaga	gaaatcatag	4260
aattttaagg	ccatcatggc	cttaatcttc	cgttccctcg	ctcactgact	cgctgcgctc	4320
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaa	gcggtaatac	ggttatccac	4380
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4440
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgcctccctg	acgagcatca	4500
caaaaatcga	cgtcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	4560
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	4620
cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	4680

```

tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 4740
gcccgaaccgc tgcgccttat ccggttaacta tcgtcttgag tccaaccggt taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 5040
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 5100
cgaaaactca cgttaaggga ttttgggtcat gagattatca aaaaggatct tcacctagat 5160
cctttttaaat taaaaatgaa gtttttaaatc aatctaaagt atatatgagt aaacttgggtc 5220
tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc tatttcgttc 5280
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 5520
atztatcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
tcaccgaggc agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac aaccaaacccg 5940
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatatct 6060
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggt gatcgagtg 6120
gtgagtaacc atgcatcac aggagtacgg ataaaatgct tgatggctcg aagaggcata 6180
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6240
ttgccatggt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcgccct cgagcaagac gtttcccggt gaatatggct cataaacacc 6420
cttgatttac tgtttatgta agcagacagt tttattgttc atgatgat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 6540
tgaagcattt atcaggggta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6718

```

&lt;210&gt; 53

&lt;211&gt; 6742

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 53

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc cataatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540

```



tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgta	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggagggtc	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgcttg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cggtgcatgg	gaacgcggat	1020
tccccgtgcc	aagagtgcgc	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcac	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggt	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttacaaaatt	cacatatata	acaacgcctg	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgcgc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcgggtgc	tgttaacggg	ggagggcag	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatategcgg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgcaccgcag	aagctgtggg	tgaccgtgta	ctacggcgtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcca	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtggtgc	tgggtgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaagg	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtag	cttcgagccc	atccccaacc	actactgcgc	ccccgcccgc	2580
ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	agggcccctg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	acttcgccga	caacgccaa	2760
gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	gcaccgccc	caacaacaac	2820
acccgcaaga	gcatccacat	cggccccggc	cgcgccttct	acaccaccgg	cgagatcatc	2880
ggcgacatcc	gccaggccca	ctgcaacctg	agccgcgcca	agtggaaacga	caccctgaac	2940
aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	ccatcgtggt	caagcacagc	3000
agcggcggcg	accccgagat	cgtgaccac	agcttcaact	gcggcggcga	gttcttctac	3060
tgcaacagca	cccagctggt	caacagcacc	tgggttcaaca	gcacctggag	caccgagggc	3120
agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	gccgcacgc	tagcagcaac	3180
aacgagagcg	agatcttccg	tccgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3240
ctgtacaagt	acaagggtgg	gaagatcgag	cccctggggc	tggccccac	caaggccaag	3300
cttaccgtcc	aggcccgcga	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	3360
cgcgccatcg	aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	3420
caggcccgcga	ccctggccgt	ggagcgcctac	ctgaaggacc	agcagcagct	cgagcagatc	3480
tggaaaccaca	ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgatc	3540
cacagcctga	tgcaggagag	ccagaaccag	cacgagaaga	acgagcagga	gctgctggag	3600
ctggacaagt	gggccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtacatc	3660
aagctgttca	tcatgatcgt	gggcggcctg	gtgggcctgc	gcacgtggtt	cgcctgctg	3720
agcatctgag	gatccagatc	tgtgtgcctt	tctagttgcc	agccatctgt	tgtttgcccc	3780
tcccccggtgc	cttcccttgac	cctggaagg	gccactccca	ctgtccttct	ctaataaaa	3840

```

gaggaaattg catcgcatthg tctgagtagg tgtcattcta ttctgggggg tgggggtgggg 3900
caggacagca aggggggagga ttgggaagac aatagcagge atgctgggga tgcgggtgggc 3960
tctatgggta cccaggtgct gaagaattga cccggttcct cctggggccag aaagaagcag 4020
gcacatcccc ttctctgtga cacaccctgt ccacgcccct ggttcttagt tccagcccca 4080
ctcataggac actcatagct caggagggtc ccgccttcaa tcccaccgcg taaagtactt 4140
ggagcgggtc ctccctccct catcagccca ccaaaccaaa cctagcctcc aagagtggga 4200
agaaattaaa gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg 4260
aggaagtaat gagagaaatc atagaatttt aaggccatca tggccttaat cttccgcttc 4320
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcgggtat cagctcactc 4380
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 4440
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 4500
gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 4560
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 4620
tccgaccctg ccgcttaccg gatacctgtc cgcttttctc ccttcgggaa gcgtggcgct 4680
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 4740
ctgtgtgcac gaaccccccg ttcagcccca ccgctgcgcc ttatccggta actatcgtct 4800
tgagtccaac ccggttaagac acgacttatc gccactggca gcagccactg gtaacaggat 4860
tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc ctaactacgg 4920
ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaag 4980
aagagttggt agctcttgat ccggcaaaac aaccaccgct ggtagcgggt gtttttttgt 5040
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 5100
tacgggggtc gacgctcagt ggaacgaaaa ctacggttaa gggatttttg tcatgagatt 5160
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 5220
aagtatatat gagtaaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 5280
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctcggggggg gggggcgctg 5340
aggtctgcct cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc 5400
cagccagaaa gtgaggagc cacggttgat gagagctttg ttgtagggtg accagttggt 5460
gattttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg 5520
atccttcaac tcagcaaaag ttcgatttat tcaacaaagc cgccgtcccg tcaagtcagc 5580
gtaatgctct gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc 5640
atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc 5700
cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg 5760
tatcggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgtca 5820
aaaataaggc tatcaagtga gaaatcacca tgagtgcaga ctgaatccgg tgagaatggc 5880
aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacg ctcgtcatca 5940
aaatcactcg catcaaccaa accgttatte attcgtgatt gcgcctgagc gagacgaaat 6000
acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac 6060
actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat 6120
gctgttttcc cggggatcgc agtggtgagt aaccatgcat catcaggagt acggataaaa 6180
tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac catctcatct 6240
gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcacgggc 6300
ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg agccatttta 6360
taccatata aatcagcatc catgttggaa tttaatcgcg gcctcgagca agacgtttcc 6420
cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt 6480
gttcgatgag atatatTTTT atcttgtgca atgtaacatc agagattttg agacacaacg 6540
tggttttccc ccccccccca ttattgaagc atttatcagg gttattgtct catgagcgga 6600
tacatatattg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 6660
aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 6720
cgtatcacga ggccctttcg tc
6742

```

&lt;210&gt; 54

&lt;211&gt; 6781

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 54

```

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggcttg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatatata acaacgcgtt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agcctgggtc ccatgcctcc 1500
agcggctcat ggtcgtctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatctcgcg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cggcaccgag aagctgtggg tgaccgtgta ctacggcgctg 2040
cccgtgtgga aggaggccac caccacctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 2160
gaggtggtgc tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaaggtgag ctctcgagccc atccccacc actactgcgc ccccgccggc 2580
ttcgccatcc tgaagtgcaa ggacaagaag ttcaacggca agggccctg caccaacgtg 2640
agcaccgtgc agtgaccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700
ggtagcctgg ccgaggagga ggtggtgatc cgcagcgcta acttcgccga caacgccaag 2760
gtgatcatcg tgcagctgaa cgagagcgct gagatcaact gcaccgccc caacaacaac 2820
acccgcaaga gcatccacat cggccccggc cgcgccttct acaccaccgg cgagatcatc 2880
ggcgacatcc gccaggcccc ctgcaacctg agccgcgcca agtggaacga caccctgaac 2940

```



aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	ccatcgtggt	caagcacagc	3000
agcggcggcg	accccgagat	cgtgaccac	agcttcaact	gcggcggcga	gttcttctac	3060
tgcaacagca	cccagctggt	caacagcacc	tggttcaaca	gcacctggag	caccgagggc	3120
agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	gccgcatcaa	gcagatcatc	3180
aacatgtggc	agaaggtggg	caaggccatg	tacgcccccc	ccatcagcgg	ccagatccgc	3240
tgcagcagca	acatcacccg	cctgctgctg	accgcgcagc	gcggaacag	caacaacgag	3300
agcgagatct	tccgtccggg	cggcggcgct	agcgccaagc	ttaccgtcca	ggcccggccg	3360
ctgctgagcg	gcatcgtgca	gcagcagaac	aacctgctgc	gcgccatcga	ggcccagcag	3420
cacctgctgc	agctgaccgt	gtggggcatc	aagcagctgc	aggcccgcac	cctggccgtg	3480
gagcgctacc	tgaaggacca	gcagcagctc	gagcagatct	ggaaccacac	cacctggatg	3540
gagtgggacc	gcgagatcaa	caactacacc	agcctgatcc	acagcctgat	cgaggagagc	3600
cagaaccagc	acgagaagaa	cgagcaggag	ctgctggagc	tggacaagtg	ggccagcctg	3660
tggaaactggt	tcaacatcac	caactggctg	tggtagatca	agctgttcat	catgatcgtg	3720
ggcggcctgg	tgggcctgcg	catcgtgttc	gccgtgctga	gcattctgagg	atccagatct	3780
gctgtgcctt	ctagttgcca	gccatctgtt	gtttgccctt	cccccgctgc	ttccttgacc	3840
ctggaaggtg	ccactcccac	tgtcctttcc	taataaaaatg	aggaaattgc	atcgcatgtg	3900
ctgagtaggt	gtcattctat	tctggggggg	ggggtggggc	aggacagcaa	gggggaggat	3960
tgggaagaca	atagcaggca	tgctggggat	gcggtgggct	ctatgggtac	ccagggtgctg	4020
aagaattgac	ccggttcctc	ctggggccaga	aagaagcagg	cacatcccct	tctctgtgac	4080
acacctgtc	cacgcccctg	gttcttagtt	ccagcccac	tcataggaca	ctcatagctc	4140
aggagggctc	cgccttcaat	cccaccgct	aaagtacttg	gagcggcttc	tccctccctc	4200
atcagcccac	caaaccaaac	ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	4260
tattaagtgc	agagggagag	aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	4320
tagaatttta	aggccatcat	ggccttaatc	ttccgcttcc	tcgctcactg	actcgctgcg	4380
ctcggtcgtt	cggctgcggc	gagcggatc	agctcactca	aaggcggtaa	tacggttatc	4440
cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	4500
gaaccgtaaa	aaggcccgct	tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	4560
tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaaccgc	acaggactat	aaagatacca	4620
ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctggt	ccgaccctgc	cgcttacccg	4680
atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	4740
gtatctcagt	tccgtgtagg	tccgtcgtc	caagctgggc	tgtgtgcacg	aacccccctg	4800
tcagcccgac	cgtgcgcct	tatccggtaa	ctatcgtctt	gagtcacacc	cggtaagaca	4860
cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	4920
cgggtgctaca	gagttcttga	agtgggtggc	taactacggc	tacactagaa	gaacagtatt	4980
tggtagctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	5040
cggcaaaaca	accaccgctg	gtagcgggtg	tttttttgtt	tgcaagcagc	agattacgcg	5100
cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acggggctctg	acgctcagtg	5160
gaacgaaaac	tcacgttaag	ggattttggt	catgagatta	tcaaaaagga	tcttcaccta	5220
gatcctttta	aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	agtaaacttg	5280
gtctgacagt	taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	gtctatttctg	5340
ttcatccata	gttgccctgac	tccggggggg	ggggcgctga	ggtctgcctc	gtgaagaagg	5400
tgttgctgac	tcataccagg	cctgaatcgc	cccatcatcc	agccagaaag	tgagggagcc	5460
acggttgatg	agagctttgt	tgtagggtga	ccagttgggtg	atcttgaaact	tttgctttgc	5520
cacggaacgg	tctgcgttgt	cgggaagatg	cgtgatctga	tccttcaact	cagcaaaagt	5580
tcgatttatt	caacaaagcc	gccgtcccgt	caagtcagcg	taatgctctg	ccagtgttac	5640
aaccaattaa	ccaattctga	ttagaaaaac	tcatcgagca	tcaaatgaaa	ctgcaattta	5700
ttcatatcag	gattatcaat	accatatttt	tgaaaaagcc	gtttctgtaa	tgaaggagaa	5760
aactcaccga	ggcagttcca	taggatggca	agatcctggt	atcggctctgc	gattccgact	5820
cgtccaacat	caatacaacc	tattaatttc	ccctcgtcaa	aaataagggt	atcaagttag	5880
aatcaccat	gagtgcgac	tgaatccggt	gagaatggca	aaagcttatg	catttctttc	5940
cagacttggt	caacaggcca	gccattacgc	tgcctcatca	aatcactcgc	atcaacaaa	6000
ccgttattca	ttcgtgattg	cgcctgagcg	agacgaaata	cgcgatcgct	gttaaaagga	6060
caattacaaa	caggaatcga	atgcaaccgg	cgcaggaaca	ctgccagcgc	atcaacaata	6120
ttttcacctg	aatcaggata	ttcttcta	acctggaatg	ctgttttccc	ggggatcgca	6180
gtggtgagta	accatgcac	atcaggagta	cggataaaat	gcttgatggt	cgggaagggc	6240

```

ataaattccg tcagccagtt tagtctgacc atctcatctg taacatcatt ggcaacgcta 6300
cctttgccat gtttcagaaa caactctggc gcatcgggct tcccatataa tcgatagatt 6360
gtcgcacctg attgcccagc attatcgcca gcccattht atccatataa atcagcatcc 6420
atgttggaat ttaatcgagg cctcgagcaa gacgtttccc gttgaatatg gtcataaca 6480
ccccctgtat tactgtttat gtaagcagac agttttattg ttcatgatga tatattttta 6540
tcttgtgcaa tgtaacatca gagattttga gacacaacgt ggctttcccc cccccccat 6600
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 6660
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 6720
gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt 6780
c 6781

```

&lt;210&gt; 55

&lt;211&gt; 6808

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 55

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc battaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500
agcggctcat ggtcgtcctg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgtctgg gctggcgctg gggcaccatg ctgctgggca 1980

```

tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	accacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tgggtgctggt	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacg	ctagcaagaa	ctgcagcttc	aacatcagca	2340
ccagcatccg	cggcaagggtg	cagaaggagt	acgccttctt	ctacaagctg	gacatcatcc	2400
ccatcgacaa	cgacaccacc	agctacagcc	tgaccagctg	caacaccagc	gtgatcacc	2460
aggcctgccc	caagggtgagc	ttcgagccca	tccccaacca	ctactgcgcc	cccgccggct	2520
tcgccatcct	gaagtgcag	gacaagaagt	tcaacggcaa	gggcccctgc	accaacgtga	2580
gcaccgtgca	gtgcacccac	ggcatccgcc	ccgtggtgag	caccagctg	ctgctgaacg	2640
gtagcctggc	cgaggaggag	gtggtgatcc	gcagcgctaa	cttcgccgac	aacgccaagg	2700
tgatcatcgt	gcagctgaac	gagagcgtgg	agatcaactg	cacccgcccc	aacaacaaca	2760
cccgcaagag	catccacatc	ggccccggcc	gcgccttcta	caccaccggc	gagatcatcg	2820
gcgacatccg	ccaggcccac	tgcaacctga	gccgcgcaa	gtggaacgac	accctgaaca	2880
agatcgtgat	caagctgcgc	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2940
gcggcggcga	ccccgagatc	gtgacccaca	gcttcaactg	cggcggcgag	ttcttctact	3000
gcaacagcac	ccagctgttc	aacagcacct	ggttcaacag	cacctggagc	accgagggca	3060
gcaacaacac	cgagggcagc	gacaccatca	ccctgccctg	ccgcatcaag	cagatcatca	3120
acatgtggca	gaagggtggc	aaggccatgt	acgccccccc	catcagcggc	cagatccgct	3180
gcagcagcaa	catcaccggc	ctgctgctga	cccgcgacgg	cggcaacagc	aacaacgaga	3240
gcgagatctt	ccgtccgggc	ggcggcgaca	tgcgcgacaa	ctggcgagc	gagctgtaca	3300
agtacaaggt	ggtgaagatc	gagcccctgg	gcgtggcccc	caccaaggcc	aagcttaccg	3360
tccaggcccc	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	ctgcgcgcca	3420
tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	ctgcaggccc	3480
gcaccctggc	cgtggagcgc	tacctgaagg	accagcagca	gctcgagcag	atctggaacc	3540
acaccacctg	gatggagtgg	gaccgcgaga	tcaacaacta	caccagcctg	atccacagcc	3600
tgatcgagga	gagccagaa	cagcaggaga	agaacgagca	ggagctgctg	gagctggaca	3660
agtgggccag	cctgtggaac	tgggtcaaca	tcaccaactg	gctgtggtac	atcaagctgt	3720
tcatcatgat	cgtggggcgg	ctggtggggc	tgcgcacgt	gttcgccgtg	ctgagcatct	3780
gaggatccag	atctgctgtg	ccttctagtt	gccagccatc	tgttgtttgc	ccctcccccg	3840
tgcttctctt	gacctggaa	ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	3900
ttgcatcgca	ttgtctgagt	aggtgtcatt	ctattctggg	gggtggggtg	gggcaggaca	3960
gcaaggggga	ggattgggaa	gacaatagca	ggcatgctgg	ggatgcggtg	ggctctatgg	4020
gtacccagggt	gctgaagaat	tgaccgggtt	cctcctgggc	cagaaagaag	caggcacatc	4080
cccttctctg	tgacacaccc	tgtccacgcc	cctggttctt	agttccagcc	ccactcatag	4140
gacactcata	gctcaggagg	gctccgcctt	caatcccacc	cgctaaagta	cttgagcgcg	4200
tctctccctc	cctcatcagc	ccaccaaacc	aaacctagcc	tccaagagtg	ggaagaaatt	4260
aaagcaagat	aggctattaa	gtgcagaggg	agagaaaatg	cctccaacat	gtgaggaagt	4320
aatgagagaa	atcatagaat	tttaaggcca	tgattttaagg	ccatcatggc	cttaatcttc	4380
cgttctctcg	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	4440
tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	4500
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	4560
ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agagggtggc	4620
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	4680
tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	4740
ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgctccaa	4800
gctgggctgt	gtgcacgaac	ccccggttca	gcccgaccgc	tgcgccttat	ccggtaacta	4860
tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	4920
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	4980
ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	5040
cggaaaaaga	gttggtagct	cttgatccgg	caaacaaacc	accgctggta	gcggtggttt	5100
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	5160
cttttctacg	gggtctgacg	ctcagtgga	cgaaaactca	cgtaaggga	ttttgggtcat	5220
gagattatca	aaaaggatct	tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	5280

```

aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 5340
acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcg gggggggggg 5400
gcgctgaggt ctgcctcgtg aagaagggtg tgctgactca taccaggcct gaatcgcccc 5460
atcatccagc cagaaagtga gggagccacg gttgatgaga gctttgttgt aggtggacca 5520
gttgggtgatt ttgaactttt gctttgccac ggaacgggtc gcgttgctcg gaagatgcgt 5580
gatctgatcc ttcaactcag caaaagtctg atttattcaa caaagccgcc gtccccgtcaa 5640
gtcagcgtaa tgctctgccg gtgtttacaac caattaacca attctgatta gaaaaactca 5700
tcgagcatca aatgaaactg caatttattc atatcaggat tatcaatacc atatttttga 5760
aaaagccggt tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga 5820
tcctgggtatc ggtctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc 5880
tcgtcaaaaa taagggtatc aagtgagaaa tcaccatgag tgacgactga atccggtgag 5940
aatggcaaaa gcttatgcat ttctttccag acttggttcaa caggccagcc attacgctcg 6000
tcatcaaaat cactcgcatc aaccaaaccc ttattcattc gtgattgcgc ctgagcgaga 6060
cgaaatacgc gatcgctggt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc 6120
aggaacactg ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc 6180
tggaatgctg ttttccccgg gatcgcatg gtgagtaacc atgcatcatc aggagtacgg 6240
ataaaatgct tgatggctcg aagaggcata aattccgtca gccagtttag tctgaccatc 6300
tcatctgtaa catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca 6360
tcgggcttcc catacaatcg atagattgtc gcacctgatt gcccgacatt atcgcgagcc 6420
catttatacc catataaatc agcatccatg ttggaattta atcgcgccct cgagcaagac 6480
gtttccccgt gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt 6540
tttattgttc atgatgatc atttttatct tgtgcaatgt aacatcagag attttgagac 6600
acaacgtggc tttccccccc ccccattat tgaagcattt atcagggtta ttgtctcatg 6660
agcgataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 6720
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 6780
aataggcgta tcacgaggcc ctttcgtc                                     6808

```

&lt;210&gt; 56

&lt;211&gt; 6769

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 56

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgct gtgtgaaata ccgcacagat gcgtgaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020

```



tccccgtgcc	aagagtgcgc	taagtaccgc	ctatagactc	tataggcaca	cccccttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatatata	acaacgcctg	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcggttac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgtggg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtgggce	accacgcct	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tggtgctggt	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcat	gatcatggag	aagggcgaga	tcaagaactg	cagcttcgct	agcaccagct	2400
gcaacaccag	cgtgatcacc	caggcctgcc	ccaaggtgag	cttcgagccc	atcccccaacc	2460
actactgcgc	ccccgcggc	ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	2520
agggcccttg	caccaacgtg	agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	2580
gcaccagct	gctgctgaac	ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	2640
acttcgccga	caacgccaa	gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	2700
gcacccgccc	caacaacaac	accgcaaga	gcatccacat	cggccccggc	cgcgccttct	2760
acaccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacctg	agccgcgcca	2820
agtggaaacga	caccctgaac	aagatcgtga	tcaagctgcg	cgagcagttc	ggcaacaaga	2880
ccatcgtgtt	caagcacagc	agcggcggcg	accccgagat	cgtgaccac	agcttcaact	2940
gcggcggcga	gttcttctac	tgcaacagca	cccagctgtt	caacagcacc	tggttcaaca	3000
gcacctggag	caccgagggc	agcaacaaca	ccgagggcag	cgacaccatc	accctgccct	3060
gccgcatcaa	gcagatcatc	aacatgtggc	agaaggtggg	caaggccatg	tacgcccccc	3120
ccatcagcgg	ccagatccgc	tgcagcagca	acatcaccgg	cctgctgctg	acccgcgacg	3180
gcggcaacag	caacaacgag	agcgagatct	tccgtccggg	cggcggcgac	atgcgcgaca	3240
actggcgcag	cgagctgtac	aagtacaagg	tgggtgaagat	cgagcccttg	ggcgtggccc	3300
ccaccaaggc	caagcttacc	gtccaggccc	gccagctgct	gagcggcatc	gtgcagcagc	3360
agaacaacct	gctgcgcgcc	atcgaggccc	agcagcacct	gctgcagctg	accgtgtggg	3420
gcatcaagca	gctgcaggcc	cgcaccctgg	ccgtggagcg	ctacctgaag	gaccagcagc	3480
agctcgagca	gatctggaac	cacaccacct	ggatggagtg	ggaccgcgag	atcaacaact	3540
acaccagcct	gatccacagc	ctgatcgagg	agagccagaa	ccagcaggag	aagaacgagc	3600
aggagctgct	ggagctggac	aagtgggcca	gcctgtggaa	ctggttcaac	atcaccaact	3660
ggctgtggta	catcaagctg	ttcatcatga	tctgtggcgg	cctggtgggc	ctgcgcacgc	3720
tgttcgccgt	gctgagcatc	tgaggatcca	gatctgctgt	gccttctagt	tgccagccat	3780
ctgttgtttg	cccccccc	gtgccttctt	tgaccctgga	aggtgccact	cccactgtcc	3840
tttctaata	aaatgaggaa	attgcacgc	attgtctgag	taggtgtcat	tctattcttg	3900
ggggtggggt	ggggcaggac	agcaaggggg	aggattggga	agacaatagc	aggcatgctg	3960
gggatgcggt	gggtctctat	ggtacccagg	tgtgaagaa	ttgaccgggt	tcctcctggg	4020
ccagaaagaa	gcaggcacat	cccccttctt	gtgacacacc	ctgtccacgc	ccctgggtct	4080
tagttccagc	cccactcata	ggacactcat	agctcaggag	ggctccgcct	tcaatcccac	4140
ccgctaaagt	acttggagcg	gtctctccct	ccctcatcag	cccaccaaac	caaacctagc	4200
ctccaagagt	gggaagaaat	taaagcaaga	taggctatta	agtgcagagg	gagagaaaat	4260
gcctccaaca	tgtgaggaag	taatgagaga	aatcatagaa	ttttaaggcc	atgatttaag	4320

```

gccatcatgg ccttaatctt ccgcttcttc gctcactgac tcgctgcgct cggtcgttcg 4380
gctgcggcga gcggtatcag ctcaactcaaa ggcggttaata cggttatcca cagaatcagg 4440
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 4500
ggccgcgctt ctggcgcttt tccataggct ccgccccctt gacgagcatc acaaaaatcg 4560
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttcccc 4620
tggaagctcc ctgctgcgct ctctgttcc gacctgccg cttaccggat acctgtccgc 4680
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggc atctcagttc 4740
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa cccccgctt agcccgaccg 4800
ctgcgcctta tccggtaact atcgtcttga gtccaaccgc gtaagacacg acttatcgcc 4860
actggcagca gccactggtg acaggattag cagagcgagg tatgtaggcg gtgctacaga 4920
gttcttgaag tgggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 4980
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac 5040
caccgctggt agcgggtggt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 5100
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtggg acgaaaactc 5160
acgttaaggg attttgggtc tgagattatc aaaaaggatc ttcacctaga tctttttaa 5220
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggc ctgacagtta 5280
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 5340
tgcctgactc gggggggggg ggcgctgagg tctgcctcgt gaagaagggt ttgctgactc 5400
ataccaggcc tgaatcgccc catcatccag ccagaaagtg agggagccac ggttgatgag 5460
agctttgttg taggtggacc agttgggtgat tttgaacttt tgctttgcca cggaacggtc 5520
tgcgttgtcg ggaagatgcg tgatctgac cttcaactca gcaaaagttc gatttattca 5580
acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtggtacaa ccaattaacc 5640
aattctgatt agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga 5700
ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcccgagg 5760
cagttccata ggatggcaag atcctggtat cggctctgca ttcgactcg tccaacatca 5820
atacaaccta ttaatttccc ctctcaaaa ataaggttat caagtggaga atcaccatga 5880
gtgacgactg aatccggtga gaatggcaaa agcttatgca tttctttcca gacttggtca 5940
acaggccagc cattacgctc gtcacaaaa tcaactcgat caaccaaacc gttattcatt 6000
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 6060
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6120
tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgcat ggtgagtaac 6180
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgtc 6240
agccagttta gtctgacctc ctcatctgta acatcattgg caacgctacc tttgccatgt 6300
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6360
tgcccgacat tatcgcgagc ccatttatac ccataataat cagcatccat gttggaattt 6420
aatcgcggcc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgtatta 6480
ctgtttatgt aagcagacag ttttattgtt catgatgata tatttttata ttgtgcaatg 6540
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 6600
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaaca 6660
ataggggttc cgcgcacatt tccccgaaa gtgccacctg acgtctaaga aaccattatt 6720
atcatgacat taacctataa aaataggcgt atcacgaggc cttttcgtc 6769

```

&lt;210&gt; 57

&lt;211&gt; 6784

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 57

```

tcgcgcgctt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120

```



```

ttggcggtg  tggggctgg  cttaactatg  cggcatcaga  gcagattgta  ctgagagtgc  180
accatatgcg  gtgtgaaata  ccgcacagat  gcgtaaggag  aaaataccgc  atcagattgg  240
ctattggcca  ttgcatacgt  tgtatccata  tcataatatg  tacatttata  ttggctcatg  300
tccaacatta  ccgccatgtt  gacattgatt  attgactagt  tattaatagt  aatcaattac  360
ggggtcatta  gttcatagcc  catatatgga  gttccgcgtt  acataactta  cggtaaattg  420
cccgcctggc  tgaccgcccc  acgacccccg  cccattgacg  tcaataatga  cgtatgttcc  480
catagtaacg  ccaataggga  ctttccattg  acgtcaatgg  gtggagtatt  tacggtaaac  540
tgcccacttg  gcagtacatc  aagtgtatca  tatgccaagt  acgcccccta  ttgacgtcaa  600
tgacggtaaa  tggcccgcc  ggcattatgc  ccagtacatg  accttatggg  actttcctac  660
ttggcagtac  atctacgtat  tagtcatcgc  tattaccatg  gtgatgcggt  tttggcagta  720
catcaatggg  cgtggatagc  ggtttgactc  acggggattt  ccaagtctcc  accccattga  780
cgtcaatggg  agtttgtttt  ggcacccaaa  tcaacgggac  ttccaaaat  gtcgtaacaa  840
ctccgcccc  ttgacgcaaa  tgggcggtag  gcgtgtacgg  tgggaggtct  atataagcag  900
agctcgttta  gtgaaccgtc  agatcgccgt  gagacgccat  ccacgtgtt  ttgacctcca  960
tagaagacac  cgggaccgat  ccagcctccg  cggccgggaa  cggtgcatlg  gaacgcggat  1020
tccccgtgcc  aagagtgcg  taagtaccgc  ctatagactc  tataggcaca  cccctttggc  1080
tcttatgcat  gctatactgt  ttttggcttg  gggcctatac  acccccgctt  ccttatgcta  1140
taggtgatgg  tatagcttag  cctatagggt  tgggttattg  accattattg  accactcccc  1200
tattggtgac  gatactttcc  attactaatc  cataacatgg  ctctttgcca  caactatctc  1260
tattggctat  atgccaatac  tctgtccttc  agagactgac  acggactctg  tatttttaca  1320
ggatggggtc  ccatttatta  tttacaaatt  cacatatata  acaacgccgt  ccccggtgcc  1380
cgcagttttt  attaaacata  gcgtgggagc  tccacgcgaa  tctcgggtac  gtgttccgga  1440
catgggctct  tctccggtag  cggcgaggct  tccacatccg  agccctgggt  ccatgcctcc  1500
agcggtcat  ggtcgctcgg  cagctccttg  ctccaaacag  tggaggccag  acttaggcac  1560
agcacaatgc  ccaccaccac  cagtgtgccc  cacaaggccg  tggcggtagg  gtatgtgtct  1620
gaaaatgagc  gtggagattg  ggctcgcacg  gctgacgcag  atggaagact  taaggcagcg  1680
gcagaagaag  atgcaggcag  ctgagttgtt  gtattctgat  aagagtcaga  ggtaactccc  1740
gttgcggtgc  tgttaacggg  ggaggcgagt  gtagtctgag  cagtactcgt  tgctgccgcg  1800
cgcgccacca  gacataatag  ctgacagact  aacagactgt  tcctttccat  gggcttttcc  1860
tgcagtcacc  gtcgtcgaca  cgtgtgatca  gatatcgccg  ccgctctaga  atgcgcgtga  1920
aggagaagta  ccagcacctg  tggcgctggg  gctggcgctg  gggcaccatg  ctgctgggca  1980
tgctgatgat  ctgcagcgcc  accgagaagc  tgtgggtgac  cgtgtactac  ggcgtgcccc  2040
tgtggaagga  ggccaccacc  accctgctct  gcgccagcga  cgccaaggcc  tacgacaccg  2100
aggtgcacaa  cgtgtgggccc  acccagcct  gcgtgcccac  cgaccccaac  cccaggagg  2160
tgggtgctgt  gaacgtgacc  gagaacttcg  acatgtggaa  gaacgacatg  gtggagcaga  2220
tgacagagga  catcatcagc  ctgtgggacc  agagcctgaa  gccctgcgtg  aagctgaccc  2280
ccctgtgctg  gagcctgaag  tgcaccgacc  tgaagaacga  caccaacacc  aacagcagca  2340
gcggccgcat  gatcatggag  aaggcgagga  tcaagaactg  cagcttcaac  atcagcacca  2400
gcacccgcgg  caaggtgcag  aaggagtacg  ccttcttcta  caagctggac  atcatcccca  2460
tcgacaacga  caccaccagc  tacagcctga  ccagctgcaa  caccagcgtg  atcaccagg  2520
cctgccccaa  ggtgagcttc  gagcccatcc  ccaaccacta  ctgcgcccc  gccggcttcg  2580
ccatcctgaa  gtgcaaggac  aagaagtcca  acggcaaggg  cccctgcacc  aacgtgagca  2640
ccgtgcagtg  caccacggc  atccgccccg  tgggtgagcac  ccagctgctg  ctgaacggta  2700
gcctggccga  ggaggagggt  gtgatccgca  gcgctaactt  cgccgacaac  gccaaaggta  2760
tcacgtgca  gctgaacgag  agcgtggaga  tcaactgcac  ccgcgctagc  gccactgca  2820
acctgagccg  cgccaagtgg  aacgacaccc  tgaacaagat  cgtgatcaag  ctgcgcgagc  2880
agttcgga  caagaccatc  gtgttcaagc  acagcagcgg  cggcgacccc  gagatcgtga  2940
cccacagctt  caactgcggc  ggcgagttct  tctactgcaa  cagcaccag  ctgttcaaca  3000
gcacctgggt  caacagcacc  tggagcaccg  agggcagcaa  caacaccgag  ggcagcgaca  3060
ccatcacctt  gccctgccgc  atcaagcaga  tcataacat  gtggcagaag  gtgggcaagg  3120
ccatgtacgc  ccccccatc  agcggccaga  tccgctgcag  cagcaacatc  accggcctgc  3180
tgctgacccg  cgacggcggc  aacagcaaca  acgagagcga  gatcttccgt  ccgggcggcg  3240
gcgacatgcg  cgacaactgg  cgcagcgagc  tgtacaagta  caaggtgggt  aagatcgagc  3300
ccctgggctg  ggccccacc  aaggccaagc  ttaccgtcca  ggcccgccag  ctgctgagcg  3360
gcacgtgca  gcagcagaac  aacctgctgc  gcgccatcga  ggcccagcag  cacctgctgc  3420

```

```

agctgaccgt gtggggcatc aagcagctgc agggccgcac cctggccgtg gagcgctacc 3480
tgaaggacca gcagcagctc gagcagatct ggaaccacac cacctggatg gagtgggacc 3540
gcgagatcaa caactacacc agcctgatcc acagcctgat cgaggagagc cagaaccagc 3600
aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg tggaaactggt 3660
tcaacatcac caactggctg tggtagatca agctgttcat catgatcgtg ggccggcctg 3720
tgggcctgcg catcgtgttc gccgtgctga gcattctgag atccagatct gctgtgcctt 3780
ctagttgcca gccatctgtt gtttgccctt ccccgctgcc ttcttgacc ctggaaggtg 3840
ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcatgtt ctgagtaggt 3900
gtcattctat tctggggggg ggggtggggc aggcagcaa gggggaggat tgggaagaca 3960
atagcaggca tgctggggat gcggtgggct ctatgggtac ccagggtgctg aagaattgac 4020
ccggttcctc ctgggccaga aagaagcagg cacatcccct tctctgtgac acaccctgtc 4080
cacgcccctg gttcttagtt ccagcccac tcataggaca ctcatagctc aggagggtc 4140
cgccttcaat cccaccgct aaagtacttg gagcgtctc tccctccctc atcagcccac 4200
caaaccaaac ctagcctcca agagtgggaa gaaattaaag caagataggtc tattaagtgc 4260
agagggagag aaaatgcctc caacatgtga ggaagtaatg agagaaatca tagaatttta 4320
aggccatgat ttaaggccat catggcctta atcttccgtt tctcgtctca ctgactcgtc 4380
gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg taatacgggt 4440
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 4500
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 4560
gcatacaaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 4620
ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac 4680
cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctctctcata gctcacgctg 4740
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 4800
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag 4860
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt 4920
aggcgggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt 4980
atltgggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 5040
atccggcaaaa caaacaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac 5100
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 5160
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 5220
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 5280
ttgggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 5340
tcgttcatcc atagttgcct gactcggggg gggggggcgc tgagggtctgc ctctgaaga 5400
agggtgtgct gactcatacc aggcctgaat cgccccatca tccagccaga aagtgaggga 5460
gccacggttg atgagagctt tggtgtaggt ggaccagttg gtgattttga acttttgctt 5520
tgccacggaa cggctctgct tgcgggaag atgcgtgate tgatccttca actcagcaaa 5580
agttcgattt attcaacaaa gccgcgctcc cgtcaagtca gcgtaatgct ctgccagtgt 5640
tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat 5700
ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 5760
gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 5820
actcgtccaa catcaataca acctattaat ttccctcgtt caaaaataag gttatcaagt 5880
gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 5940
ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcataacc 6000
aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa 6060
ggacaattac aaacaggaat cgaatgcaac cggcgcagga aactgcccag cgcataaca 6120
atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc 6180
gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcggaaga 6240
ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 6300
ctacctttgc catgtttcag aaacaactct ggcgcacggt gcttcccata caatcgatag 6360
attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata taaatcagca 6420
tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat atggctcata 6480
acacccttg tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt 6540
ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc ccccccccc 6600
cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6660
tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 6720

```

taagaaacca ttattatcat gacattaacc tataaaaata ggcgatcac gaggcccttt 6780  
cgtc 6784

<210> 58

<211> 6787

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 58

tgcgcgcttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
ccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480  
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600  
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660  
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960  
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcgt cccccgtgcc 1380  
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
agcggctcat ggtcgtctcg cagctccttg ctccataacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920  
aggagaagta ccagcacctg tggcgtggtg gctggcgctg gggcaccatg ctgctgggca 1980  
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040  
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100  
aggtgcacaa cgtgtgggac acccagcct gcgtgcccac cgaccccaac cccagaggag 2160  
tgggtgctgt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220  
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280  
ccctgtgctg gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340  
gcggccgcgt gatcatggag aagggcgaga tcaagaactg cagcttcaac atcagcacca 2400  
gcatccgcgg caaggtgcag aaggagtacg ccttcttcta caagctggac atcatcccca 2460



tcgacaacga	caccaccage	tacagcctga	ccagctgcaa	caccagcgtg	atcaccacagg	2520
cctgccccaa	ggtgagcttc	gagcccatcc	ccaaccacta	ctgcgcccc	gccggcttcg	2580
ccatcctgaa	gtgcaaggac	aagaagttca	acggcaaggg	cccctgcacc	aacgtgagca	2640
ccgtgcagtg	caccacaggc	atccgccccg	tggtgagcac	ccagctgctg	ctgaacggta	2700
gcctggccga	ggaggagggtg	gtgatccgca	gcgctaactt	cgccgacaac	gccaagggtga	2760
tcatacgtgca	gctgaacgag	agcgtggaga	tcaactgcac	ccgccccaac	aacaacaccc	2820
gcaagagcat	ccacatcggc	cccggccgcg	ccttctacac	caccggcgag	atcatcggcg	2880
acatccgcca	ggcccactgc	aacctgagcc	gcgccaagtg	gaacgacacc	ctgaacaaga	2940
tcgtgatcaa	gctgcgcgag	cagttcggca	acaagaccat	cgtgttcaag	cacagcagcg	3000
gcggcgaccc	cgagatcgtg	accacagct	tcaactgcgg	cggcgagttc	ttctactgca	3060
acagcgctag	cctgccctgc	cgcatacaagc	agatcatcaa	catgtggcag	aaggtgggca	3120
aggccatgta	cgcccccccc	atcagcggcc	agatccgctg	cagcagcaac	atcaccggcc	3180
tgctgctgac	ccgcgacggc	ggcaacagca	acaacgagag	cgagatcttc	cgtccgggcg	3240
gcggcgacat	gcgcgacaac	tggcgagcag	agctgtacaa	gtacaagggtg	gtgaagatcg	3300
agccccctggg	cgtggccccc	accaaggcca	agcttaccgt	ccaggcccg	cagctgctga	3360
gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggccag	cagcacctgc	3420
tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccg	caccctggcc	gtggagcgct	3480
acctgaagga	ccagcagcag	ctcgagcaga	tctggaacca	caccacctgg	atggagtggg	3540
accgcgagat	caacaactac	accagcctga	tccacagcct	gatcgaggag	agccagaacc	3600
agcaggagaa	gaacgagcag	gagctgctgg	agctggacaa	gtgggccagc	ctgtggaact	3660
ggttcaacat	caccaactgg	ctgtggtaca	tcaagctgtt	catcatgatc	gtgggcgggc	3720
tggtgggcct	gcgcatacgt	ttcgccgtgc	tgagcatctg	aggatccaga	tctgctgtgc	3780
cttctagtgt	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	accctggaag	3840
gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcatacgc	tgtctgagta	3900
ggtgtcattc	tattctgggg	ggtgggggtg	ggcaggacag	caagggggag	gattgggaag	3960
acaatagcag	gcatgctggg	gatgcgggtg	gctctatggg	taccaggtg	ctgaagaatt	4020
gacccggttc	ctcctggggc	agaaagaagc	aggcacatcc	ccttctctgt	gacacaccct	4080
gtccacgccc	ctggttctta	gttccagccc	cactcatagg	acactcatag	ctcaggaggg	4140
ctccgccttc	aatcccaccc	gctaaagtac	ttggagcgg	ctctccctcc	ctcatcagcc	4200
caccaaacca	aacctagcct	ccaagagtgg	gaagaaatta	aagcaagata	ggctattaag	4260
tgcagagggg	gagaaaatgc	ctccaacatg	tgaggaagta	atgagagaaa	tcatagaatt	4320
ttaaggccat	gatttaaggc	catcatggcc	ttaatcttcc	gcttctctgc	tactgactc	4380
gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	4440
gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	gccagcaaaa	4500
ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	4560
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	4620
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgcgct	4680
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	4740
ctgtagggtat	ctcagttcgg	tgtaggctgt	tgcctccaag	ctgggctgtg	tgcacgaacc	4800
ccccgttcag	cccgaccgct	gcgccttate	cggtaactat	cgtcttgagt	ccaacccgg	4860
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	4920
tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	4980
agtatttggt	atctgcgctc	tgtgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	5040
ttgatccggc	aaacaaacca	ccgctggtag	cgggtggttt	tttgtttgca	agcagcagat	5100
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgate	ttttctacgg	ggtctgacgc	5160
tcagtggaa	gaaaactcac	gttaagggtat	tttggctcat	agattatcaa	aaaggatctt	5220
cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	5280
aacttgggtc	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	5340
atttcgttca	tccatagttg	cctgactcgg	gggggggggg	cgtgagggtc	tgcctcgtga	5400
agaagggtgt	gctgactcat	accaggcctg	aatcgcccca	tcattccagcc	agaaagtgag	5460
ggagccacgg	ttgatgagag	ctttgttgta	ggtggaccag	ttggtgattt	tgaacttttg	5520
ctttgccacg	gaacgggtctg	cgttgtcggg	aagatgcgtg	atctgatect	tcaactcagc	5580
aaaagttcga	tttattcaac	aaagccgccc	tcccgtcaag	tcagcgtaat	gctctgccag	5640
tgtttacaacc	aattaaccaa	ttctgattag	aaaaactcat	cgagcatcaa	atgaaactgc	5700
aattttattca	tatcaggatt	atcaatacca	tatttttgaa	aaagccgttt	ctgtaatgaa	5760

```

ggagaaaact caccgaggca gttccatagg atggcaagat cctgggtatcg gtctgcgatt 5820
ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aaggttatca 5880
agtgagaaat caccatgagt gacgactgaa tccgggtgaga atggcaaaaag cttatgcatt 5940
tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcatca 6000
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatcacgc atcgctgtta 6060
aaaggacaat tacaacacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 6120
acaatatatt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6180
atcgcgatgg tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6240
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6300
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6360
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6420
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6480
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6540
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tttccccccc 6600
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6660
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6720
gtctaagaaa ccattattat catgacatta acctataaaa ataggcggtat caccgagccc 6780
tttcgtc 6787

```

&lt;210&gt; 59

&lt;211&gt; 6691

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 59

```

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcg ctagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg acctatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcgg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttatbg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500

```

```

agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcgggtg tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatacgcgg cgcgtctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccagcct gcgtgcccac cgacccaac cccaggagg 2160
tgggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaggcctg cccaagggtg agcttcgagc ccatcccaa ccactactgc gccccgcgcg 2400
gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caaggggccc tgcaccaacg 2460
tgagcacctg gcagtgacc cagggcatcc gcccgtggt gagcaccag ctgctgctga 2520
acggtagcct ggccgaggag gaggtggtga tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgc cccaacaaca 2640
acaccgcaa gagcatccac atcgccccg gccgcgctt ctacaccacc ggcgagatca 2700
tcggcgacat ccgccaggcc cactgcaacc tgagccgcgc caagtggaa gacaccctga 2760
acaagatcgt gatcaagctg cgcgagcagt tcggcaacaa gaccatcgtg ttcaagcaca 2820
gcagcggcgg cgaccccgag atcgtgaccc acagcttcaa ctgcggcggc gagttcttct 2880
actgcaacag caccagctg ttcaacagca cctggttcaa cagcacctgg agcaccgagg 2940
gcagcaacaa caccgagggc agcgacacca tcaccctgcc ctgcgcgcatc aagcagatca 3000
tcaacatgtg gcagaagggt ggcaaggcca tgtacgccc cccatcagc ggccagatcc 3060
gctgcagcag caacatcacc ggctgctgc tgaccgcga cggcggcaac agcaacaacg 3120
agagcgagat ctccgctccg ggccggcggc acatgcgcga caactggcgc agcgagctgt 3180
acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc cccaccaag gccaaagctta 3240
ccgtccaggc ccgccagctg ctgagcggca tcgtgcagca gcagaacaac ctgctgcgcg 3300
ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg 3360
cccgcaccct ggccgtggag cgctacctga aggaccagca gcagctcgag cagatctgga 3420
accacaccac ctggatggag tgggaccgcg agatcaacaa ctacaccagc ctgatccaca 3480
gcctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg ctggagctgg 3540
acaagtgggc cagcctgtgg aactggttca acatcaccaa ctggtgtgtg tacatcaagc 3600
tgttcatcat gatcgtgggc ggctggtgg gcctgcgcac cgtgttcgcc gtgctgagca 3660
tctgaggatc cagatctgct gtgccttcta gttgccagcc atctgttgtt tgccccctcc 3720
ccgtgccttc cttgaccctg gaaggtgcca ctccactgt cctttcctaa taaaatgagg 3780
aaattgcac gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagg 3840
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 3900
tgggtaccca ggtgctgaag aattgacccg gttcctcctg ggccagaaag aagcaggcac 3960
atccccctct ctgtgacaca cctgtccac gcccctgggt cttagtcca gccccactca 4020
taggacactc atagctcagg agggctccgc cttcaatccc acccgctaaa gtacttgagg 4080
cggctctctc ctccctcatc agcccaccaa accaaaccta gcctccaaga gtgggaagaa 4140
attaaagcaa gataggctat taagtgcaga gggagagaaa atgcctcaa catgtgagga 4200
agtaatgaga gaaatcatag aattttaagg ccatgatbta aggccatcat ggccttaatc 4260
ttccgcttcc tcgctcactg actcgctgcg ctgggtcgtt cggctgcggc gagcggtatc 4320
agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa 4380
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 4440
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg 4500
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 4560
ctctcctgtt ccgaccctgc cgtttaccgg atacctgtcc gcctttctcc cttcgggaag 4620
cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcgggtgtagg tcgttcgctc 4680
caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgctt tatccggtaa 4740
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 4800

```



```

taacaggatt agcagagcga ggtatgtagg cgggtgctaca gagttcttga agtgggtggcc 4860
taactacggc tacactagaa gaacagtatt tggatatctgc gctctgctga agccagttac 4920
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcgggtg 4980
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5040
gatcttttct acgggggtctg acgctcagtg gaacgaaaac tcacgttaag ggatttttgt 5100
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 5160
atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 5220
ggcacctatc tcagcgatct gtctatttctg ttcattccata gttgcctgac tcgggggggg 5280
ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 5340
cccatcatcc agccagaaag tgaggagacc acggttgatg agagctttgt tbtagggtga 5400
ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg 5460
cgtgatctga tctttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccgt 5520
caagtcagcg taatgctctg ccagtggtac aaccaattaa ccaattctga ttagaaaaac 5580
tcacgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt 5640
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 5700
agatcctggt atcgggtctgc gattccgact cgtccaacat caatacaacc tattaatttc 5760
ccctcgtcaa aaataagggt atcaagttag aaatcaccat gagtgacgac tgaatccggt 5820
gagaatggca aaagcttatg catttcttct cagacttggt caacaggcca gccattacgc 5880
tcgtcatcaa aatcactcgc atcaaccaa ccgttattca ttcgtgattg cgctgagcg 5940
agacgaaata cgcgatcgtg gttaaaagga caattacaaa caggaatcga atgcaaccgg 6000
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaat 6060
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcac atcaggagta 6120
cggataaaat gcttgatggt cggaagaggg ataaattccg tcagccagtt tagtctgacc 6180
atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc 6240
gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgc attatcgca 6300
gcccatttat acccatataa atcagcatcc atggttggat ttaatcgcg cctcgagcaa 6360
gacgtttccc gttgaatatg gtcataaca ccccttgat tactgtttat gtaagcagac 6420
agttttattg ttcattgatg tatattttta tcttggtgca tgtaacatca gagattttga 6480
gacacaacgt ggctttcccc ccccccccat tattgaagca tttatcaggg ttattgtctc 6540
atgagcggat acataatttg atgtatttag aaaaataaac aaataggggt tccgcgcaca 6600
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 6660
aaaaataggc gtatcacgag gccctttcgt c 6691

```

&lt;210&gt; 60

&lt;211&gt; 6727

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 60

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg gccattatgc ccagtacatg accttatggg actttcctac 660

```

```

ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatth gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatgggggtc ccatttatta tttacaaatt cacatatata acaacgcggt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggctcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
agggtcacaa cgtgtggggc acccagcct gcgtgccac cgaccccaac ccccaggagg 2160
tgggtgctgg gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaagggt cagaaggagt acgccttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 2460
aggcctgccc caaggtgagc ttcgagccca tccccacca ctactgcgc cccgcccggct 2520
tcgccatcct gaagtgaag gacaagaagt tcaacggcaa gggcccctgc accaacgtga 2580
gcaccgtgca gtgcaccac ggcattccgc ccgtgggtgag caccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgcccag aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcgct agcgcctact 2760
gcaacctgag ccgcgcgaag tggaaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2820
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2880
tgaccacag cttcaactgc ggcggcgagt tcttctactg caacagcacc cagctgttca 2940
acagcacctg gttcaacagc acctggagca ccgagggcag caacaacacc gagggcagcg 3000
acaccatcac cctgccctgc cgcattcaagc agatcatcaa catgtggcag aaggtgggca 3060
aggccatgta ccccccccc atcagcggcc agatccgctg cagcagcaac atcaccggcc 3120
tgctgctgac ccgcgacggc ggcaacagca acaacgagag cgagatcttc cgtccgggcg 3180
gcggcgacat gcgcgacaac tggcgcgagc agctgtacaa gtacaagggt gtgaagatcg 3240
agcccctggg cgtggcccc accaaggcca agcttaccgt ccaggcccgc cagctgctga 3300
gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag cagcacctgc 3360
tgcagctgac cgtgtggggc atcaagcagc tgcaggccc caccctggcc gtggagcgct 3420
acctgaagga ccagcagcag ctgcagcaga tctggaacca caccacctgg atggagtggg 3480
accgcgagat caacaactac accagcctga tccacagcct gatcgaggag agccagaacc 3540
agcaggagaa gaacgagcag gagctgctgg agctggacaa gtgggcccag ctgtgggaact 3600
ggttcaacat caccaactgg ctgtggtaca tcaagctgtt catcatgatc gtgggcccgc 3660
tgggtgggct gcgcacgtg ttcgccgtgc tgagcatctg aggatccaga tctgctgtgc 3720
cttctagttg ccagccatct gttgtttgcc cctccccgt gccttccttg accctggaag 3780
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcac tgtctgagta 3840
ggtgtcatte tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag 3900
acaatagcag gcatgctggg gatgcgggtg gctctatggg taccaggtg ctgaagaatt 3960

```

gacccgggttc ctccctggggc agaaagaagc aggcacatcc ccttctctgt gacacaccc 4020  
gtccacgccc ctgggttctta gttccagccc cactcatagg acactcatag ctcaggagg 4080  
ctccgccttc aatcccaccc gctaaagtac ttggagcggc ctctccctcc ctcacagcc 4140  
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4200  
tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa tcatagaatt 4260  
ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc tcaactgactc 4320  
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4380  
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4440  
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4500  
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag 4560  
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4620  
taccggatac ctgtccgctt ttctcccttc gggagcgtg gcgctttctc atagctcacg 4680  
ctgtaggtat ctcagttcgg tgtaggctgt tcgctccaag ctgggctgtg tgcacgaacc 4740  
ccccgttcag ccgcaccgct gcgccttata cggtaactat cgtcttgagt ccaaccgggt 4800  
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4860  
tgtaggcggc gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4920  
agtatttggc atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 4980  
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 5040  
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgaecg 5100  
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5160  
cacctagatc cttttaaatt aaaaatgaag ttttaaatac atctaaagta tatatgagta 5220  
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5280  
atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc tgcctcgtga 5340  
agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagttag 5400  
ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5460  
ctttgccacg gaacgggtctg cgttgctcggg aagatgcgtg atctgatcct tcaactcagc 5520  
aaaagttcga tttattcaac aaagccgccg tcccgtaag tcagcgtaat gctctgccag 5580  
tggtacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5640  
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5700  
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5760  
ccgactcgtc caacatcaat acaacctatt aatttccctc cgtcaaaaat aaggttatca 5820  
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5880  
tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcatac 5940  
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 6000  
aaaggacaat tacaacacag aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 6060  
acaatatatt cacctgaatc aggatattct tctaatacct ggaatgctgt tttccggggg 6120  
atcgcagtgg tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6180  
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6240  
acgctacctt tgccatgttt cagaaacaac tctggcgcct cgggcttccc atacaatcga 6300  
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6360  
gcatccatgt tggaatttaa tcgctggcct gagcaagacg tttcccgttg aatatggctc 6420  
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6480  
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tttccccccc 6540  
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6600  
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6660  
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgat cagaggccc 6720  
tttcgctc 6727

&lt;210&gt; 61

&lt;211&gt; 6730

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 61

```

tcgcgcggtt  cggatgatgac  ggtgaaaacc  tctgacacat  gcagctcccg  gagacggtca  60
cagcttgtct  gtaagcggat  gccgggagca  gacaagcccc  tcagggcgcg  tcagcgggtg  120
ttggcgggtg  tcggggctgg  cttaactatg  cggcatcaga  gcagattgta  ctgagagtgc  180
accatatgcg  gtgtgaaata  ccgcacagat  gcgtaaggag  aaaataccgc  atcagattgg  240
ctattggcca  ttgcatacgt  tgtatccata  tcataatatg  tacatttata  ttggctcatg  300
tccaacatta  ccgccatgtt  gacattgatt  attgactagt  tattaatagt  aatcaattac  360
ggggtcatta  gttcatagcc  catatatgga  gttccgcgtt  acataactta  cggtaaatgg  420
cccgcctggc  tgaccgcca  acgacccccg  cccattgacg  tcaataatga  cgtatgttcc  480
catagtaacg  ccaataggga  ctttccattg  acgtcaatgg  gtggagtatt  tacggtaaac  540
tgcccacttg  gcagtacatc  aagtgtatca  tatgccaaag  acgcccccta  ttgacgtcaa  600
tgacggtaaa  tggcccgctt  ggcattatgc  ccagtacatg  accttatggg  actttcctac  660
ttggcagtac  atctacgtat  tagtcacgc  tattaccatg  gtgatgcggt  tttggcagta  720
catcaatggg  cgtggatagc  ggtttgactc  acggggattt  ccaagtctcc  accccattga  780
cgtcaatggg  agtttgtttt  ggcacaaaaa  tcaacgggac  ttccaaaat  gtcgtaacaa  840
ctccgcccc  ttgacgcaa  tgggcggtag  gcgtgtacgg  tgggaggtct  atataagcag  900
agctcgttta  gtgaaccgtc  agatcgcttg  gagacgccat  ccacgctgtt  ttgacctcca  960
tagaagacac  cgggaccgat  ccagcctccg  cggccgggaa  cgggtgcattg  gaacgcggat  1020
tccccgtgcc  aagagtgcg  taagtaccgc  ctatagactc  tataggcaca  cccctttggc  1080
tcttatgcat  gctatactgt  ttttggttg  gggcctatac  acccccgctt  ccttatgcta  1140
taggtgatgg  tatagcttag  cctataggtg  tgggttattg  accattattg  accactcccc  1200
tattggtgac  gatactttcc  attactaatc  cataacatgg  ctctttgcca  caactatctc  1260
tattggctat  atgccaatac  tctgtccttc  agagactgac  acggactctg  tatttttaca  1320
ggatggggtc  ccattttatta  ttacaaaatt  cacatataca  acaacgccgt  ccccgctgcc  1380
cgcagttttt  attaaacata  gcgtgggata  tccacgcgaa  tctcgggtac  gtgttcggga  1440
catgggctct  tctccggtag  cggcggagct  tccacatccg  agccctggtc  ccatgcctcc  1500
agcggctcat  ggtcgtcgg  cagctccttg  ctctaacag  tggaggccag  acttaggcac  1560
agcacaatgc  ccaccaccac  cagtgtgccg  cacaaggccg  tggcggtagg  gtatgtgtct  1620
gaaaatgagc  gtggagattg  ggctcgcacg  gctgacgcag  atggaagact  taaggcagcg  1680
gcagaagaag  atgcaggcag  ctgagttgtt  gtattctgat  aagagtcaga  ggtaactccc  1740
gttgcggtgc  tgttaacggg  ggagggcagt  gtagtctgag  cagtactcgt  tgctgcgcgc  1800
cgcgccacca  gacataatag  ctgacagact  aacagactgt  tcttttccat  gggctttttc  1860
tgcagtcacc  gtcgtcgaca  cgtgtgatca  gatatcgcg  ccgctctaga  atgcgcgtga  1920
aggagaagta  ccagcacctg  tggcgctggg  gctggcgctg  gggcaccatg  ctgctgggca  1980
tgctgatgat  ctgcagcgcc  accgagaagc  tgtgggtgac  cgtgtactac  ggcgtgcccg  2040
tgtggaagga  ggccaccacc  accctgctct  gcgccagcga  cgccaaggcc  tacgacaccg  2100
aggtgcacaa  cgtgtgggac  acccagcct  gcgtgccac  cgacccaac  cccaggagg  2160
tgggtgctgg  gaacgtgacc  gagaacttcg  acatgtggaa  gaacgacatg  gtggagcaga  2220
tgcaagagga  catcatcagc  ctgtgggacc  agagcctgaa  gccctgcgtg  aagctgaccc  2280
ccctgtgcgt  gagcctgaag  tgcaccgacg  ctagcaagaa  ctgcagcttc  aacatcagca  2340
ccagcatccg  cggcaagggt  cagaaggagt  acgccttctt  ctacaagctg  gacatcatcc  2400
ccatcgacaa  cgacaccacc  agctacagcc  tgaccagctg  caacaccagc  gtgatcacc  2460
aggcctgccc  caaggtgagc  ttcgagccca  tccccaacca  ctactgcgcc  cccgcgggt  2520
tcgccatcct  gaagtgcaag  gacaagaagt  tcaacggcaa  gggcccctgc  accaacgtga  2580
gcaccgtgca  gtgcaccac  ggcacccgcc  ccgtgggtgag  caccagctg  ctgctgaacg  2640
gtagcctggc  cgaggaggag  gtggtgatcc  gcagcgctaa  ctctgcgcgac  aacgccaagg  2700
tgatcatcgt  gcagctgaac  gagagcgtgg  agatcaactg  caccgcgcc  aacaacaaca  2760
cccgaagag  catccacatc  ggccccggcc  gcgccttcta  caccaccggc  gagatcatcg  2820
gcgacatccg  ccaggccacc  tgcaacctga  gccgcgccaa  gtggaacgac  accctgaaca  2880
agatcgtgat  caagctgcgc  gagcagttcg  gcaacaagac  catcgtgttc  aagcacagca  2940
gcggcggcga  ccccgagatc  gtgaccaca  gcttcaactg  cggcggcgag  ttcttctact  3000
gcaacagcgc  tagcctgccc  tgccgcatca  agcagatcat  caacatgtgg  cagaagggtg  3060

```



```

gcaaggccat gtacgcccc cccatcagcg gccagatccg ctgcagcagc aacatcaccg 3120
gcctgctgct gacccgcgac ggcggaaca gcaacaacga gagcgagatc ttccgtccgg 3180
gcggcgggcg catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtgaaga 3240
tcgagccctt gggcgtggcc cccaccaagg ccaagcttac cgtccaggcc cgccagctgc 3300
tgagcggcat cgtgcagcag cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc 3360
tgctgcagct gaccgtgtgg ggcatcaagc agctgcaggc ccgcaccctg gccgtggagc 3420
gctacctgaa ggaccagcag cagctcgagc agatctggaa ccacaccacc tggatggagt 3480
gggaccgcga gatcaacaac tacaccagcc tgatccacag cctgatcgag gagagccaga 3540
accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc agcctgtgga 3600
actggttcaa catcaccaac tggctgtggt acatcaagct gttcatcatg atcgtgggcg 3660
gcctggtggg cctgcgcctc gtgttcgccc tgctgagcat ctgaggatcc agatctgctg 3720
tgcccttctag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg 3780
aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg cattgtctga 3840
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 3900
aagacaatag caggcatgct ggggatgcgg tgggctctat gggtaaccag gtgctgaaga 3960
attgaccggg ttctcctggg gccagaaaga agcaggcaca tccccttctc tgtgacacac 4020
cctgtccacg cccctgggtc ttagttccag cccactcat aggacactca tagctcagga 4080
gggctccgcc ttcaatccca cccgctaaag tacttgagc ggtctctccc tccctcatca 4140
gcccaccaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt 4200
aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag aaatcataga 4260
atthtaaggc catgatttaa ggccatcatg gccttaatct tccgcttctc cgctcactga 4320
ctcgtgcgc tcggtcgttc ggctgcggcg agcgggtatca gctcactcaa aggcggtaab 4380
acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca 4440
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccataggc tccgcccccc 4500
tgacgagcat cacaaaaatc gacgtcaag tcagaggtgg cgaaaccgca caggactata 4560
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 4620
gcttacggga tactgtccg cttttctccc ttccgggaagc gtggcgcttt ctcatagctc 4680
acgctgtagg tatctcagtt cgggtgtagg cgttcgctcc aagctgggct gtgtgcacga 4740
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaacc 4800
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4860
gtatgtaggc ggtgctacag agttcttgaa gtgggtggct aactacggct acactagaag 4920
aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttggtag 4980
ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgg ttttttgttt gcaagcagca 5040
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5100
cgctcagtgg aacgaaaact cacgttaagg gattttgggt atgagattat caaaaaggat 5160
cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga 5220
gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5280
tctatttcgt tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5340
tgaagaagg gtgtgctgact cataccagc ctgaatcgcc ccacatcca gccagaaagt 5400
gagggagcca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5460
ttgctttgcc acggaacggt ctgcgttgtc ggggaagatgc gtgatctgat ccttcaactc 5520
agcaaaagt cgatttatcc aacaaagccg cgtcccgtc aagtcagcgt aatgctctgc 5580
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5640
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5700
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcgggtctgcg 5760
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataaggtta 5820
tcaagtgaga aatcaccatg agtgacgact gaatccgggt agaatggcaa aagcttatgc 5880
atthctttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 5940
tcaaccaaac cgttatccat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 6000
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6060
tcaacaatat ttccacctga atcaggatat tcttctaata cctggaatgc tgttttccc 6120
gggatcgagc tgggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6180
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6240
gcaacgctac ctttgccatg ttccagaaac aactctggcg catcgggctt cccatacaat 6300
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6360

```

```

tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6420
ctcataacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6480
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6540
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6600
tgtattttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6660
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6720
ccctttcgtc                                     6730

```

&lt;210&gt; 62

&lt;211&gt; 6688

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 62

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcaagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccc 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccagcct gcgtgcccac cgacccaac cccagaggag 2160

```



tggtgctggt	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtggggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcat	gatcatggag	aagggcgaga	tcaagaactg	cagcttcgct	agcaccagct	2400
gcaacaccag	cgtgatcacc	caggcctgcc	ccaaggtgag	cttcgagccc	atccccaacc	2460
actactgcgc	ccccgccggc	ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	2520
agggcccttg	caccaacgtg	agcaccgtgc	agtgcaccca	cggcatccgc	cccgtgggtga	2580
gcaccagct	gctgctgaac	ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	2640
acttcgccga	caacgccaa	gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	2700
gcaccgcgc	tagcgccac	tgcaacctga	gccgcgcaa	gtggaacgac	accctgaaca	2760
agatcgtgat	caagctgcgc	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2820
gcggcggcga	ccccgagatc	gtgacccaca	gcttcaactg	cggcggcgag	ttcttctact	2880
gcaacagcac	ccagctgttc	aacagcacct	ggttcaacag	cacctggagc	accgagggca	2940
gcaacaacac	cgagggcagc	gacaccatca	ccctgccctg	ccgcatcaag	cagatcatca	3000
acatgtggca	gaaggtgggc	aaggccatgt	acgccccccc	catcagcggc	cagatccgct	3060
gcagcagcaa	catcaccggc	ctgctgctga	cccgcgacgg	cggcaacagc	aacaacgaga	3120
gcgagatctt	ccgtccgggc	ggcggcgaca	tgcgcgacaa	ctggcgcagc	gagctgtaca	3180
agtacaaggt	ggtgaagatc	gagcccttgg	gcgtggcccc	caccaaggcc	aagcttaccg	3240
tccaggcccc	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	ctgcgcgcca	3300
tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	ctgcaggccc	3360
gcaccctggc	cgtggagcgc	tacctgaagg	accagcagca	gctcgagcag	atctggaacc	3420
acaccacctg	gatggagtgg	gaccgcgaga	tcaacaacta	caccagcctg	atccacagcc	3480
tgatcgagga	gagccagaac	cagcaggaga	agaacgagca	ggagctgctg	gagctggaca	3540
agtgggccag	cctgtggaac	tggttcaaca	tcaccaactg	gctgtggtac	atcaagctgt	3600
tcacatgat	cgtggggcgc	ctggtggggc	tgcgcacgt	gttcgccgtg	ctgagcatct	3660
gaggatccag	atctgctgtg	ccttctagtt	gccagccatc	tgttgtttgc	ccctcccccg	3720
tgccttcctt	gaccctggaa	ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	3780
ttgcatcgca	ttgtctgagt	aggtgtcatt	ctattctggg	gggtggggtg	gggcaggaca	3840
gcaaggggga	ggattgggaa	gacaatagca	ggcatgctgg	ggatgcgggtg	ggctctatgg	3900
gtacccaggt	gctgaagaat	tgacccggtt	cctcctgggc	cagaaagaag	caggcacatc	3960
cccttctctg	tgacacaccc	tgtccacgcc	cctggttcct	agttccagcc	ccactcatag	4020
gacactcata	gctcaggagg	gctccgcctt	caatcccacc	cgctaaagta	cttgagcgcg	4080
tctctccctc	cctcatcagc	ccaccaaacc	aaacctagcc	tccaagagtg	ggaagaaatt	4140
aaagcaagat	aggctattaa	gtgcagaggg	agagaaaatg	cctccaacat	gtgagggaat	4200
aatgagagaa	atcatagaat	tttaaggcca	tgatttaagg	ccatcatggc	cttaatcttc	4260
cgcttctctg	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	4320
tcactcaaag	gcggtaatat	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	4380
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	4440
ccataggctc	cgcccccttg	acgagcatca	caaaaatcga	cgctcaagtc	agaggtggcg	4500
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	togtgcgctc	4560
tctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	4620
ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgctccaa	4680
gctgggctgt	gtgcacgaac	ccccggttca	gcccgaccgc	tgcgccttat	ccggtaacta	4740
tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	4800
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	4860
ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	4920
cggaaaaaga	gttggttagct	cctgatccgg	caaacaaacc	accgctggta	gcggtgggtt	4980
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	5040
cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgtttaaggga	ttttgggtcat	5100
gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	gttttaaatc	5160
aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	5220
acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcg	gggggggggg	5280
gcgctgaggt	ctgcctcgtg	aagaagggtg	tgctgactca	taccaggcct	gaatcgcccc	5340
atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	gctttgttgt	aggtggacca	5400
gttgggtgatt	ttgaactttt	gctttgccac	ggaacgggtc	gcgttgtcgg	gaagatgcgt	5460

```

gatctgatcc ttcaactcag caaaagttcg atttattcaa caaagccgcc gtcccgtcaa 5520
gtcagcgtaa tgctctgcca gtgttacaac caattaacca attctgatta gaaaaactca 5580
tcgagcatca aatgaaactg caattttattc atatcaggat tatcaatacc atatttttga 5640
aaaagccggt tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga 5700
tcctgggtatc ggtctgcatc tccgactcgt ccaacatcaa tacaacctat taatttcccc 5760
tcgtcaaaaa taagggtatc aagtgagaaa tcaccatgag tgacgactga atccgggtgag 5820
aatggcaaaa gcttatgcat ttctttccag acttggtcaa caggccagcc attacgctcg 5880
tcatcaaaat cactcgcac aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga 5940
cgaaatacgc gatcgtctgtt aaaaggacaa tcacaaacag gaatcgaatg caaccggcgc 6000
aggaacactg ccagcgcac aacaatattt tcacctgaat caggatattc ttctaatacc 6060
tggaatgctg ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg 6120
ataaaatgct tgatgggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc 6180
tcatctgtaa catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca 6240
tcgggcttcc catacaatcg atagattgtc gcacctgatt gcccgcacatt atcgcgagcc 6300
catttatacc catataaatc agcatccatg ttggaattta atcgcggcct cgagcaagac 6360
gtttcccgtt gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt 6420
tttattgttc atgatgatat atttttatct tgtgcaatgt aacatcagag attttgagac 6480
acaacgtggc tttccccccc ccccatattt tgaagcattt atcagggtta ttgtctcatg 6540
agcggataca tatttgaatg tathtagaaa aataaacaaa taggggttcc gcgcacattt 6600
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 6660
aataggcgta tcacgaggcc ctttcgtc 6688

```

&lt;210&gt; 63

&lt;211&gt; 6691

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 63

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320

```

```

ggatgggggtc ccatttatta tttaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgtggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccagcct gcgtgccccc cgacccaac cccaggagg 2160
tgggtgctggg gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgacgagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgacct 2280
ccctgtgctg gagcctgaag tgcaccgacc tgaagaacga caccaacacc aacagcagca 2340
gcggccgcat gatcatggag aagggcgaga tcaagaactg cagcttcgct agcaccagct 2400
gcaacaccag cgtgatcacc caggcctgcc ccaagggtgag cttcgagccc atccccaacc 2460
actactgccc ccccgccggc ttccgcatcc tgaagtgcaa ggacaagaag ttcaacggca 2520
agggcccctg caccaacgtg agcaccgtgc agtgaccca cggcatccgc cccgtgggtga 2580
gcacccagct gctgctgaac ggtagcctgg ccgaggagga ggtgggtgatc cgcagcgcta 2640
acttcgccga caacgccaa gtgatcatcg tgcagctgaa cgagagcgtg gagatcaact 2700
gcacccgccc caacaacaac acccgcaaga gcatccacat cggccccggc cgcgccttct 2760
acaccaccgg cgagatcatc ggcgacatcc gccaggccca ctgcaacctg agccgcgcca 2820
agtggaacga caccctgaac aagatcgtga tcaagctgcg cgagcagttc ggcaacaaga 2880
ccatcgtgtt caagcacagc agcggcggcg accccgagat cgtgaccac agcttcaact 2940
gcggcggcga gttcttctac tgcaacagcg cttagcctgcc ctgcgcatac aagcagatca 3000
tcaacatgtg gcagaagggt ggcaaggcca tgtacgcccc ccccatcagc ggccagatcc 3060
gctgcagcag caacatcacc ggcctgctgc tgacccgcga cggcggcaac agcaacaacg 3120
agagcgagat cttccgtccg ggcggcggcg acatgcgcga caactggcgc agcagagctgt 3180
acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc cccaccaag gccaaagctta 3240
ccgtccaggc ccgccagctg ctgagcggca tctgagcaga gcagaacaac ctgctgcgcg 3300
ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg 3360
cccgaccct ggccgtggag cgctacctga aggaccagca gcagctcgag cagatctgga 3420
accacaccac ctggatggag tgggaccgcg agatcaacaa ctacaccagc ctgatccaca 3480
gcctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg ctggagctgg 3540
acaagtgggc cagcctgtgg aactgggtca acatcaccaa ctggctgtgg tacatcaagc 3600
tgttcatcat gatcgtgggc ggcctggtgg gcctgcgcac cgtgttcgcc gtgctgagca 3660
tctgaggatc cagatctgct gtgccttcta gttgccagcc atctgttgtt tgccccctcc 3720
ccgtgccttc cttgaccctg gaaggtgcc a ctcccactgt cctttcctaa taaaatgagg 3780
aaattgcata gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagg 3840
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 3900
tgggtacca ggtgctgaag aattgaccgg gttcctcctg ggccagaaag aagcaggcac 3960
atccccctct ctgtgacaca cctgtccac gccctgggt cttagtcca gccccactca 4020
taggacactc atagctcagg agggctccgc cttcaatccc acccgctaaa gtacttgagg 4080
cggctctctc ctccctcatc agcccaccaa accaaacct gctccaaga gtgggaagaa 4140
attaaagcaa gataggctat taagtgcaga gggagagaaa atgcctccaa catgtgagga 4200
agtaatgaga gaaatcatag aattttaagg ccatgatbta aggccatcat ggccttaate 4260
ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggatc 4320
agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa 4380
catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgct 4440
tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 4500
gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg 4560
ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 4620

```



```

cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 4680
caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa 4740
ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 4800
taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 4860
taactacggc tacactagaa gaacagtatt tggatatctgc gctctgctga agccagttac 4920
cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcgggtg 4980
tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 5040
gatcttttct acgggggtctg acgctcagtg gaacgaaaac tcacgttaag ggatttttgg 5100
catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttta 5160
atcaatctaa agtatatatg agtaaaactg gtctgacagt taccaatgct taatcagtga 5220
ggcacctatc tcagcgatct gtctatttcg ttcattccata gttgcctgac tcgggggggg 5280
ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg cctgaatcgc 5340
cccatcatcc agccagaaaag tgaggagacc acggttgatg agagctttgt tgtagggtga 5400
ccagttgggt attttgaact tttgctttgc cacggaacgg tctgcgttgt cggaagatg 5460
cgtgatctga tccttcaact cagcaaaaagt tcgatttatt caacaaagcc gccgtcccgt 5520
caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac 5580
tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatatt 5640
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca 5700
agatcctggc atcgggtctg gattccgact cgtccaacat caatacaacc tattaatttc 5760
ccctcgtcaa aaataagggt atcaagttag aaatcaccat gagtgcgcac tgaatccgg 5820
gagaatggca aaagcttatg catttctttc cagacttggt caacaggcca gccattacgc 5880
tcgtcatcaa aatcactcgc atcaaccaa cgtttattca ttcgtgattg cgctgagcg 5940
agacgaaata cgcgatcgtc gttaaaagga caattacaaa caggaatcga atgcaaccgg 6000
cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttcta 6060
acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcac atcaggagta 6120
cggataaaat gcttgatggc cggaagaggg ataaattccg tcagccagtt tagtctgacc 6180
atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc 6240
gcacggggct tcccatacaa tcgatagatt gtcgcacctg attgcccgc attatcgca 6300
gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcg cctcgagcaa 6360
gacgtttccc gttgaatatg gtcataaca ccccttgat tactgtttat gtaagcagac 6420
agttttattg ttcattgatg tatattttta tcttgtagaa tgtaacatca gagattttga 6480
gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg ttattgtctc 6540
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 6600
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 6660
aaaaataggg gtatcacgag gccctttcgt c 6691

```

&lt;210&gt; 64

&lt;211&gt; 6706

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 64

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgcattcata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480

```

catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgc	tattaccatg	gtgatgcgg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgcttg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
ggttgcggtc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgttgcgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgcctgg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgtgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgccc	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	accacgcctt	gcgtgcccac	cgaccccaac	ccccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcgcccgcat	gatcatggag	aagggcgaga	tcaagaactg	cagcttcaac	atcagcacca	2400
gcatccgcgg	caaggtgcag	aaggagtaag	ccttcttcta	caagctggac	atcatcccca	2460
tgcacaacga	caccaccagc	tacagcctga	ccagctgcaa	caccagcgtg	atcaccaggg	2520
cctgccccaa	ggtgagcttc	gagcccatcc	ccaaccacta	ctgcgcccc	gccggcttcg	2580
ccatcctgaa	gtgcaaggac	aagaagttca	acggcaaggg	cccctgcacc	aacgtgagca	2640
ccgtgcagtg	caccacgggc	atccgccccg	tgggtgagcac	ccagctgctg	ctgaacggta	2700
gcctggccga	ggaggagggt	gtgatccgca	gcgctaactt	cgcgcacaac	gccaagggtga	2760
tcatcgtgca	gctgaacgag	agcgtggaga	tcaactgcac	ccgcgctagc	gcccactgca	2820
acctgagccg	cgcgaagtgg	aacgacaccc	tgaacaagat	cgtgatcaag	ctgcgcgagc	2880
agttcggcaa	caagaccatc	gtgttcaagc	acagcagcgg	cggcgacccc	gagatcgtga	2940
cccacagctt	caactgcggc	ggcgagttct	tctactgcaa	cagcgttagc	ctgccctgcc	3000
gcatcaagca	gatcatcaac	atgtggcaga	aggtgggcaa	ggccatgtac	gcccccccca	3060
tcagcggcca	gatccgctgc	agcagcaaca	tcaccggcct	gctgctgacc	cgcgacggcg	3120
gcaacagcaa	caacgagagc	gagatcttcc	gtccggggcg	cggcgacatg	cgcgacaact	3180
ggcgcagcga	gctgtacaag	tacaagggtg	tgaagatcga	gcccctgggc	gtggccccca	3240
ccaaggccaa	gcttaccgtc	caggccccgc	agctgctgag	cggcatcgtg	cagcagcaga	3300
acaacctgct	gcgcgccatc	gaggcccagc	agcacctgct	gcagctgacc	gtgtggggca	3360
tcaagcagct	gcaggccccg	accctggccg	tggagcgcta	cctgaaggac	cagcagcagc	3420
tgcagcagat	ctggaaccac	accacctgga	tggagtggga	ccgcgagatc	aacaactaca	3480
ccagcctgat	ccacagcctg	atcgaggaga	gccagaacca	gcaggagaag	aacgagcagg	3540
agctgctgga	gctggacaag	tgggccagcc	tgtggaactg	gttcaacatc	accaactggc	3600
tgtgggtacat	caagctgttc	atcatgatcg	tgggcggcct	gggtggcctg	cgcacgtgtg	3660
tgcgcgtgct	gagcatctga	ggatccagat	ctgctgtgcc	ttctagtgtg	cagccatctg	3720
ttgtttgccc	ctcccccgct	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	3780

```

cctaataaaa tgaggaaatt gcatcgcat gtctgagtag gtgtcattct attctggggg 3840
gtgggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg 3900
atgcggtggg ctctatgggt acccaggtgc tgaagaattg acccggttcc tcctggggcca 3960
gaaagaagca ggcacatccc cttctctgtg acacaccctg tccacgcccc tggttcttag 4020
ttccagcccc actcatagga cactcatagc tcaggagggc tccgccttca atcccacccg 4080
ctaaagtact tggagcggtc tctccctccc tcatcagccc accaaaccaa acctagcctc 4140
caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag agaaaatgcc 4200
tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg atttaaggcc 4260
atcatggcct taatcttccg cttcctcgct cactgactcg ctgctcgg tcttccggct 4320
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4380
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4440
cgcggttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 4500
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 4560
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 4620
tctcccttcg ggaagcgtgg cgcttttcca tagctcacgc tgtaggtatc tcagttcggg 4680
gtaggtcggt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg 4740
cgcttatacc ggtaactatc gtcttgagtc caaccgggta agacacgact tatcgccact 4800
ggcagcagcc actggtaaca ggattagcag agcgagggtat gtaggcgggtg ctacagagtt 4860
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct 4920
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 4980
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 5040
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg 5100
ttaagggatt ttgggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 5160
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttgggtctg acagttacca 5220
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5280
ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaagggtgtg ctgactcata 5340
ccaggcctga atcgcccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5400
tttggtgtag gtggaccagt tgggtgattt gaacttttgc tttgccacgg aacgggtctgc 5460
gttgctcgga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5520
aagccgctg cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5580
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5640
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5700
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5760
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5820
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5880
ggccagccat tacgctcgtc atcaaaatca ctgcgcatcaa ccaaaccggt attcattcgt 5940
gattgcgcct gagcgagacg aaatacgcga tgcgtgttaa aaggacaatt acaaacagga 6000
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 6060
ggatattctt ctaatacctg gaatgctgtt ttcccgggga tgcgagtggg gagtaaccat 6120
gcatcatcag gactacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6180
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6240
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6300
ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaat 6360
cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccctt tgtattactg 6420
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6480
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 6540
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6600
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6660
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6706

```

&lt;210&gt; 65

&lt;211&gt; 6649

&lt;212&gt; DNA



&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 65

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg ggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga atgcgcgta 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgccc 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccagcctt gcgtgcccac cgaccccaac cccaggagg 2160
tgggtgctgg gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgctg aagctgaccc 2280
ccctgtgctg gagcctgaag tgcaccgacg ctagcaagaa ctgcagcttc aacatcagca 2340
ccagcatccg cggcaagggt cagaaggagt acgcttctt ctacaagctg gacatcatcc 2400
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 2460
aggcctgccc caaggtgagc ttcgagccca tccccacca ctactgcgc cccgcggct 2520
tcgccatcct gaagtgaag gacaagaagt tcaacggcaa gggcccctgc accaacgtga 2580
gcaccgtgca gtgcaccac ggcattccgc ccgtgggtgag caccagctg ctgctgaacg 2640
gtagcctggc cgaggaggag gtggtgatcc gcagcgctaa cttcgcccac aacgccaagg 2700
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcgct agcgcctact 2760
gcaacctgag ccgcgccaag tggaaacgaca ccctgaacaa gatcgtgatc aagctgcgcg 2820
agcagttcgg caacaagacc atcgtgttca agcacagcag cggcggcgac cccgagatcg 2880
tgaccacag cttcaactgc ggcggcgagt tcttctactg caacagcgt agcctgcct 2940

```

```

gccgcatcaa gcagatcatc aacatgtggc agaaggtggg caaggccatg tacgcccccc 3000
ccatcagcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg acccgcgacg 3060
gcggaacag caacaacgag agcgagatct tccgtccggg cggcggcgac atgcgcgaca 3120
actggcgag cgagctgtac aagtacaagg tggatgaagat cgagcccttg ggcgtggccc 3180
ccaccaaggc caagcttacc gtccaggccc gccagctgct gagcggcatac gtgcagcagc 3240
agaacaacct gctgcgcgcc atcgaggccc agcagcacct gctgcagctg accgtgtggg 3300
gcatcaagca gctgcaggcc cgcaccctgg ccgtggagcg ctacctgaag gaccagcagc 3360
agctcgagca gatctggaac cacaccacct ggatggagtg ggaccgagag atcaacaact 3420
acaccagcct gatccacagc ctgatcgagg agagccagaa ccagcaggag aagaacgagc 3480
aggagctgct ggagctggac aagtgggcca gcctgtggaa ctggttcaac atcaccaact 3540
ggctgtggta catcaagctg ttcattcatga tgcgtggcgg cctgggtgggc ctgcgcatac 3600
tgttcgccgt gctgagcatc tgaggatcca gatctgctgt gccttctagt tgccagccat 3660
ctgttgtttg cccctccccg gtgccttctc tgaccctgga aggtgccact cccactgtcc 3720
tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattcttg 3780
ggggtggggg ggggcaggac agcaaggggg aggattggga agacaatagc aggcagctg 3840
gggatgcggg gggctctatg ggtaccagg tgctgaagaa ttgaccgggt tcctcctggg 3900
ccagaaagaa gcaggcacat ccccttctct gtgacacacc ctgtccacgc ccctggttct 3960
tagttccagc cccactcata ggacactcat agctcaggag ggctccgcct tcaatccac 4020
ccgctaaagt acttgagcgc gtctctccct cctcatcag cccaccaaac caaacctagc 4080
ctccaagagt ggaagaaat taaagcaaga taggctatta agtgcagagg gagagaaaat 4140
gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc atgatttaag 4200
gccatcatgg ccttaatctt ccgcttctc gctcactgac tcgctgcgct cggtcgttcg 4260
gctgcggcga gcggtatcag ctactcaaaa ggcggttaata cggttatcca cagaatcagg 4320
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 4380
ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc aaaaaaatcg 4440
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttcccc 4500
tggaagctcc ctgctgcgct ctctgttcc gaccctgccg cttaccggat acctgtccgc 4560
ctttctccct tcgggaagcg tggcgcttct tcatagctca cgtgtagggt atctcagttc 4620
gggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg 4680
ctgcgcctta tccggtaact atcgtcttga gtccaaccgc gtaagacacg acttatcgcc 4740
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 4800
gttcttgaag tgggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc 4860
tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac 4920
caccgctggg agcgggtgggt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 4980
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 5040
acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga tcctttttaa 5100
ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggg ctgacagtta 5160
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 5220
tgctgactc gggggggggg ggcgctgagg tctgcctcgt gaagaagggtg ttgctgactc 5280
ataccaggcc tgaatcgccc catcatccag ccagaaagtg agggagccac ggttgatgag 5340
agctttgttg taggtggacc agttgggtgat tttgaacttt tgctttgcca cggaaagggtc 5400
tgcgttgtcg ggaagatgcg tgatctgac cttcaactca gcaaaagttc gatttattca 5460
acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtggtacaa ccaattaacc 5520
aattctgatt agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga 5580
ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctaccgagg 5640
cagttccata ggatggcaag atcctgggat cggctctgca ttcgactcg tccaacatca 5700
atacaacctt ttaatttccc ctctcaaaa ataagggtat caagtgagaa atcaccatga 5760
gtgacgactg aatccgggtga gaatggcaaa agcttatgca tttctttcca gacttgttca 5820
acaggccagc cattacgctc gtcatacaaa tcaactcgat caaccaaac gttattcatt 5880
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 5940
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6000
tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgaggt ggtgagtaac 6060
catgcatcat caggagtacg gataaaatgc ttgatggctg gaagaggcat aaattccgtc 6120
agccagttta gtctgacct ctcatctgta acatcattgg caacgctacc tttgccatgt 6180
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6240

```

```

tgcccgacat tatcgcgagc ccattttatac ccatataaat cagcatccat gttggaattt 6300
aatcgcgggc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgtatta 6360
ctgtttatgt aagcagacag ttttattggt catgatgata tattttttatc ttgtgcaatg 6420
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 6480
tatcaggggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacia 6540
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6600
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc 6649

```

&lt;210&gt; 66

&lt;211&gt; 6610

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 66

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacia 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgaggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgcccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctcgggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgcgc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgbtaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatecgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccc 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggce acccacgcct gcgtgcccac cgaccccaac cccagaggag 2160

```



tggtgctggt	gaacgtgacc	gagaacttctg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgctgt	aagctgacce	2280
ccctgtgctg	gagcctgaag	tgcaccgacc	tgaagaacga	caccaacacc	aacagcagca	2340
gcggccgcat	gatcatggag	aagggcgaga	tcaagaactg	cagcttcgct	agcaccagct	2400
gcaacaccag	cgtgatcacc	caggcctgcc	ccaaggtgag	cttcgagccc	atccccaacc	2460
actactgctg	ccccgcccgc	ttcgccatcc	tgaagtgcaa	ggacaagaag	ttcaacggca	2520
agggccccctg	caccaacgtg	agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	2580
gcacccagct	gctgctgaac	ggtagcctgg	ccgaggagga	ggtggtgatc	cgcagcgcta	2640
acttcgcccga	caacgcccaag	gtgatcatcg	tgcagctgaa	cgagagcgtg	gagatcaact	2700
gcacccgctg	tagcgcccac	tgaacctga	gccgcgcaa	gtggaacgac	accctgaaca	2760
agatcgtgat	caagctgctg	gagcagttcg	gcaacaagac	catcgtgttc	aagcacagca	2820
gcggcgccga	ccccgagatc	gtgacccaca	gcttcaactg	cggcgccgag	ttcttctact	2880
gcaacagctg	tagcctgccc	tgcgcgcatc	agcagatcat	caacatgtgg	cagaaggtgg	2940
gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	ctgcagcagc	aacatcaccg	3000
gcctgctgct	gacccgctgac	ggcggaaca	gcaacaacga	gagcgagatc	ttccgtccgg	3060
gcggcgccga	catgcgctgac	aactggcgca	gcgagctgta	caagtacaag	gtggtgaaga	3120
tcgagcccc	gggctgtggc	cccaccaagg	ccaagcttac	cgtccaggcc	cgccagctgc	3180
tgagcgccat	cgtgcagcag	cagaacaacc	tgctgcgctg	catcgaggcc	cagcagcacc	3240
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcaggc	ccgcaccctg	gccgtggagc	3300
gctacctgaa	ggaccagcag	cagctcgagc	agatctggaa	ccacaccacc	tggatggagt	3360
gggaccgctg	gatcaacaac	tacaccagcc	tgatccacag	cctgatcgag	gagagccaga	3420
accagcagga	gaagaacgag	caggagctgc	tggagctgga	caagtgggcc	agcctgtgga	3480
actggttcaa	catcaccaac	tggctgtggt	acatcaagct	gttcatcatg	atcgtgggctg	3540
gcctggtggg	cctgcgcatc	gtgttcgctg	tgctgagcat	ctgaggatcc	agatctgctg	3600
tgccttctag	ttgccagcca	tctgttgttt	gccccctccc	cgtgccttcc	ttgaccctgg	3660
aaggtgccac	tcccactgtc	ctttccta	aaaatgagga	aattgcatcg	cattgtctga	3720
gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	gaggattggg	3780
aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	gggtaccag	gtgctgaaga	3840
attgaccggg	ttcctcctgg	gccagaaaga	agcaggcaca	tccccctctc	tgtgacacac	3900
cctgtccacg	cccctgggtc	ttagttccag	ccccactcat	aggacactca	tagctcagga	3960
gggctccgcc	ttcaatccca	cccgttaaag	tacttggagc	ggtctctccc	ttcctcatca	4020
gcccacaaa	ccaaacctag	cctccaagag	tgggaagaaa	ttaaagcaag	ataggctatt	4080
aagtgcagag	ggagagaaaa	tgcctccaac	atgtgaggaa	gtaatgagag	aaatcataga	4140
atthtaaggc	catgatttaa	ggccatcatg	gccttaatct	tccgcttctc	cgctcactga	4200
ctcgctgctg	tcgggtcggtc	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	4260
acgggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	4320
aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcggtt	ttccataggc	ttccgcccc	4380
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga	caggactata	4440
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	4500
gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	4560
acgctgtagg	tatctcagtt	cggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	4620
accccccggt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaacc	4680
ggtaagacac	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	4740
gtatgtaggc	ggtgctacag	agttcttgaa	gtgggtggcct	aactacggct	acactagaag	4800
aacagtattt	ggtatctgctg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	4860
ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	4920
gattacgctg	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	4980
cgctcagtg	aacgaaaact	cacgttaagg	gatttttggtc	atgagattat	caaaaaggat	5040
cttcacctag	atccttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	gtatatatga	5100
gtaaaacttg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	cagcgatctg	5160
tctatttcgt	tcatccatag	ttgcctgact	cggggggggg	gggcgctgag	gtctgcctcg	5220
tgaagaagg	ggtgctgact	cataccaggc	ctgaatcgcc	ccatcatcca	gccagaaagt	5280
gagggagcca	cgggtgatga	gagctttgtt	gtaggtggac	cagttggtga	ttttgaactt	5340
ttgctttgcc	acggaacggg	ctgcgttgtc	gggaagatgc	gtgatctgat	ccttcaactc	5400
agcaaaagtt	cgatttatcc	aacaaagccg	ccgtcccgtc	aagtcagcgt	aatgctctgc	5460

```

cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5520
tgcaatttat tcatatcagg attatcaata ccataatttt gaaaaagccg tttctgtaat 5580
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 5640
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataagggtta 5700
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5760
atttctttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 5820
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 5880
ttaaaggagac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 5940
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttccc 6000
gggatcgagc tgggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6060
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6120
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6180
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6240
tcagcatcca tgggtggaatt taatcgcggc ctcgagcaag acgtttccc tgaatatgg 6300
ctcataacac cccttgatb actgtttatg taagcagaca gttttattgt tcatgatgat 6360
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6420
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6480
tgtattttaga aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6540
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6600
cccttctgtc                                     6610

```

&lt;210&gt; 67

&lt;211&gt; 6610

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 67

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcgg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttagcgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380

```

```

cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggte ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgetgcccgc 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatctcgcg cgcctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgetgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgccagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtgggccc acccacgcct gcgtgcccac cgaccccaac cccagaggag 2160
tgggtgctggt gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgcacgagga catcatcagc ctgtgggacc agagcctgaa gccctgctg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaggcctg ccccaagggtg agcttcgagc ccatcccca cactactgc gccccgcgcg 2400
gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caagggcccc tgcaccaacg 2460
tgagcacctg gcagtgaccc cacggcatcc gcccgtggt gagcaccag ctgctgctga 2520
acggtagcct ggccgaggag gaggtggtga tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgcg gctagcgcgc 2640
actgcaacct gagccgcgcc aagtggaaag acaccctgaa caagatcgtg atcaagctgc 2700
gcgagcagtt cggcaacaag accatcgtgt tcaagcacag cagcggcggc gaccccgaga 2760
tcgtgacca cagcttcaac tgcggcggcg agttcttcta ctgcaacagc acccagctgt 2820
tcaacagcac ctggttcaac agcacctgga gcaccgagg cagcaacaac accgagggca 2880
gcgacaccat caccctgccc tgccgcacatc agcagatcat caacatgtgg cagaagggtg 2940
gcaaggccat gtacgcccc cccatcagcg gccagatccg ctgcagcagc aacatcaccg 3000
gcctgctgct gacccgcgac ggccgcaaca gcaacaacga gagcgagatc ttccgtccgg 3060
gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtgaaga 3120
tcgagccctt gggcgtggcc cccaccaagg ccaagcttac cgtccaggcc cgccagctgc 3180
tgagcggcat cgtgcagcag cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc 3240
tgctgcagct gaccgtgtgg ggcacaaagc agctgcaggg ccgcaccctg gccgtggagc 3300
gctacctgaa ggaccagcag cagctcgagc agatctggaa ccacaccacc tggatggagt 3360
gggaccgcga gatcaacaac tacaccagcc tgatccacag cctgatcgag gagagccaga 3420
accagcagga gaagaacgag caggagctgc tggagctgga caagtgggccc agcctgtgga 3480
actggttcaa catcaccaac tggctgtggt acatcaagct gttcatcatg atcgtgggcg 3540
gcctggtggg cctgcgcacg gtgttcgcgc tgetgagcat ctgaggatcc agatctgctg 3600
tgcttcttag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgacctgg 3660
aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg cattgtctga 3720
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 3780
aagacaatag caggcatgct ggggatgcgg tgggctctat gggtaaccag gtgctgaaga 3840
attgacccgg ttctcctcgg gccagaaaga agcaggcaca tccccttctc tgtgacacac 3900
cctgtccacg cccctggttc ttagttccag cccactcat aggacactca tagctcagga 3960
gggtccgcgc ttcaatecca cccgctaaag tacttgagc ggtctctccc tccctcatca 4020
gcccaccaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt 4080
aagtgcagag ggagagaaaa tgccctcaac atgtgaggaa gtaatgagag aaatcataga 4140
atlttaaggc catgatttaa ggccatcatg gccttaatct tccgcttcc cgtcactga 4200
ctcgtgcgc tcggtcgttc ggctgcggcg agcggatca gctcactcaa aggcggtaat 4260
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 4320
aaaggccagg aaccgtaaaa aggcgcgctt gctggcgttt tcccataggc tccgcccccc 4380
tgacgagcat caaaaaatc gacgtcaag tcagaggtgg cgaaacccga caggactata 4440
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc 4500
gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt ctcatagctc 4560
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 4620
acccccggt cagcccagac gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 4680

```



```

ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4740
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 4800
aacagtatct ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 4860
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgg ttttttggtt gcaagcagca 4920
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cgggggtctga 4980
cgctcagtg gacgaaaact caggttaagg gattttggtc atgagattat caaaaaggat 5040
cttcacctag atcctttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga 5100
gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5160
tctatttcgt tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5220
tgaagaagg gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5280
gagggagcca cgggtgatga gagctttgtt gtaggtggac cagttgggtga ttttgaactt 5340
ttgctttgcc acggaacggc ctgcgttgct ggggaagatgc gtgatctgat ccttcaactc 5400
agcaaaagtt cgattttatt aacaaagccg cgcgtccgct aagtcagcgt aatgctctgc 5460
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5520
tgcaatttat tcatatcagg attatcaata ccataatctt gaaaaagccg tttctgtaat 5580
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcgggtctgcg 5640
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataagggtta 5700
tcaagtgaga aatcaccatg agtgacgact gaatccgggt agaatggcaa aagcttatgc 5760
atttctttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 5820
tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 5880
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 5940
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6000
gggatcgag tggtagtaaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6060
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6120
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6180
cgatagattg tcgcacctga ttgcccgaac ttatcgcgag cccatttata cccatataaa 6240
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6300
ctcataaac cccttgatc actgtttatg taagcagaca gttttattgt tcatgatgat 6360
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6420
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6480
tgtattttag aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6540
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6600
ccctttcgtc 6610

```

&lt;210&gt; 68

&lt;211&gt; 6613

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 68

```

tcgcgcggtt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcggggtg tcgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600

```

tgacggtaaa	tggcccgccct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgc	tattaccatg	gtgatgcgg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggatbt	ccaagtctcc	acccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggctt	atataagcag	900
agctcgttta	gtgaaccgct	agatcgccct	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccccttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	accccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggagc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	atgcgcgtga	1920
aggagaagta	ccagcacctg	tggcgctggg	gctggcgctg	gggcaccatg	ctgctgggca	1980
tgctgatgat	ctgcagcgcc	accgagaagc	tgtgggtgac	cgtgtactac	ggcgtgcccg	2040
tgtggaagga	ggccaccacc	accctgctct	gcgccagcga	cgccaaggcc	tacgacaccg	2100
aggtgcacaa	cgtgtggggc	accacgcct	gcgtgccac	cgacccaac	ccccaggagg	2160
tgggtgctgg	gaacgtgacc	gagaacttcg	acatgtggaa	gaacgacatg	gtggagcaga	2220
tgcacgagga	catcatcagc	ctgtgggacc	agagcctgaa	gccctgcgtg	aagctgaccc	2280
ccctgtgcgt	gagcctgaag	tgcaccgacg	ctagcaccag	ctgcaacacc	agcgtgatca	2340
cccaggcctg	ccccagggtg	agcttcgagc	ccatcccca	ccactactgc	gccccgcgcg	2400
gcttcgccat	cctgaagtgc	aaggacaaga	agttcaacgg	caaggggccc	tgcaccaacg	2460
tgagcacctg	gcagtgcacc	cacggcatcc	gccccgtgg	gagcacccag	ctgctgctga	2520
acggtagcct	ggccgaggag	gaggtgggtga	tccgcagcgc	taacttcgcc	gacaacgcca	2580
aggtgatcat	cgtgcagctg	aacgagagcg	tggagatcaa	ctgcaccgcg	cccaacaaca	2640
acacccgcaa	gagcatccac	atcgggcccc	gccgcgcctt	ctacaccacc	ggcgagatca	2700
tcggcgacat	ccgccaggcc	cactgcaacc	tgagccgcgc	caagtggaa	gacacctga	2760
acaagatcgt	gatcaagctg	cgcgagcagt	tcggcaacaa	gaccatcgtg	ttcaagcaca	2820
gcagcggcgg	cgaccccgag	atcgtgaccc	acagcttcaa	ctgcggcggc	gagttcttct	2880
actgcaacag	cgctagcctg	ccctgcccga	tcaagcagat	catcaacatg	tggcagaagg	2940
tgggcaaggc	catgtacgcc	ccccccatca	gcggccagat	ccgctgcagc	agcaacatca	3000
ccggcctgct	gctgaccgcg	gacggcggca	acagcaacaa	cgagagcgag	atcttccgtc	3060
cgggcggcgg	cgacatgcgc	gacaactggc	gcagcgagct	gtacaagtac	aagggtgggtga	3120
agatcgagcc	cctgggcgtg	gccccacca	aggccaagct	taccgtccag	gcccgcgcgc	3180
tgctgagcgg	catcgtgcag	cagcagaaca	acctgctgcg	cgccatcgag	gcccagcagc	3240
acctgctgca	gctgaccgtg	tggggcatca	agcagctgca	ggcccgacc	ctggccgtgg	3300
agcgtacct	gaaggaccag	cagcagctcg	agcagatctg	gaaccacacc	acctggatgg	3360
agtgggaccg	cgagatcaac	aactacacca	gcctgatcca	cagcctgate	gaggagagcc	3420
agaaccagca	ggagaagaac	gagcaggagc	tgctggagct	ggacaagtgg	gccagcctgt	3480
ggaactgggt	caacatcacc	aactggctgt	ggtacatcaa	gctgttcac	atgatcgtgg	3540
gcggcctgg	gggcctgcgc	atcgtgttcg	ccgtgctgag	catctgagga	tccagatctg	3600
ctgtgccttc	tagttgccag	ccatctgttg	tttgcccttc	ccccgtgcct	tccttgacct	3660
tggaagggtg	cactcccact	gtcctttcct	aataaaatga	ggaaattgca	tcgcattgtc	3720
tgagtaggtg	tcattctatt	ctgggggggtg	gggtggggca	ggacagcaag	ggggaggatt	3780
gggaagacaa	tagcaggcat	gctgggggatg	cgggtgggctc	tatgggtacc	cagggtgctga	3840
agaattgacc	cgggttcctcc	tggggccagaa	agaagcaggc	acatcccttc	ctctgtgaca	3900

caccctgtcc acgcccctgg ttcttagtbc cagcccact cataggacac tcatagctca 3960  
ggagggctcc gccttcaatc ccaccgcta aagtacttgg agcgggtctc ccctccctca 4020  
tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc aagataggct 4080  
attaagtga gagggagaga aaatgcctcc aacatgtgag gaagtaatga gagaaatcat 4140  
agaattttaa ggccatgatt taaggccatc atggccttaa tcttccgctt cctcgcctcac 4200  
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggc 4260  
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 4320  
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc 4380  
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 4440  
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 4500  
gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag 4560  
ctcacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca 4620  
cgaaccccc gttcagcccc accgctgcgc cttatccggc aactatcgtc ttgagtccaa 4680  
cccggtaaga cagcacttat cgccactggc agcagccact ggtaacagga ttagcagagc 4740  
gaggtatgta ggcgggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag 4800  
aagaacagta tttgggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 4860  
tagctcttga tccggcaaac aaaccaccgc tggtagcggc ggtttttttg tttgcaagca 4920  
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacgggggc 4980  
tgacgctcag tggaaacgaaa actcacgtta agggattttg gtcattgagat tatcaaaaag 5040  
gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 5100  
tgagtaaact tgggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 5160  
ctgtctatct cgttcatcca tagttgcctg actcgggggg gggggggcgt gaggtctgcc 5220  
tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc gcccacatcat ccagccagaa 5280  
agtgaggag ccacgggtga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa 5340  
cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa 5400  
ctcagcaaaa gttcgattta ttcaacaaag ccgcgcgtcc gtcaagtcag cgtaatgctc 5460  
tgccagtggt acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga 5520  
aactgcaatt tattcatatc aggtattatc ataccatatt tttgaaaaag ccgtttctgt 5580  
aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcgggtc 5640  
gcgattccga ctctgccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg 5700  
ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagctta 5760  
tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc 5820  
gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg 5880  
ctgttaaaaag gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc 5940  
gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc 6000  
ccgggggatcg cagtgggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg 6060  
gtcgggaagag gcataaatc cgtcagccag tttagtctga ccatctcatc tgtaacatca 6120  
ttggcaacgc tacctttgcc atgtttcaga acaactctg gcgcacggg cttcccatac 6180  
aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccatct ataccatata 6240  
aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc ccgttgaata 6300  
tggctcataa cacccttgt attactgttt atgtaagcag acagttttat tgttcatgat 6360  
gatataatctt tatcttgtgc aatgtaacat cagagatctt gagacacaac gtggctttcc 6420  
ccccccccc attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt 6480  
gaatgtatct agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc 6540  
cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg 6600  
aggcccttcc gtc 6613

&lt;210&gt; 69

&lt;211&gt; 6532

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;



<223> plasmid pVR1012x/s containing HIV genes

<400> 69

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtag atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgcat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatatata acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagc gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga atgcgcgtga 1920
aggagaagta ccagcacctg tggcgctggg gctggcgctg gggcaccatg ctgctgggca 1980
tgctgatgat ctgcagcgcc accgagaagc tgtgggtgac cgtgtactac ggcgtgcccg 2040
tgtggaagga ggccaccacc accctgctct gcgcagcga cgccaaggcc tacgacaccg 2100
aggtgcacaa cgtgtggggc acccagcct gcgtgcccac cgaccccaac cccagggagg 2160
tggtgctggg gaacgtgacc gagaacttcg acatgtggaa gaacgacatg gtggagcaga 2220
tgacagagga catcatcagc ctgtgggacc agagcctgaa gccctgcgtg aagctgaccc 2280
ccctgtgcgt gagcctgaag tgcaccgacg ctagcaccag ctgcaacacc agcgtgatca 2340
cccaggcctg cccaagggtg agcttcgagc ccaccccaa ccaactactg cccccgccc 2400
gcttcgccat cctgaagtgc aaggacaaga agttcaacgg caagggcccc tgcaccaacg 2460
tgagcaccgt gcagtgcacc cacggcatcc gcccgtggg gagcaccag ctgctgctga 2520
acggtagcct ggccgaggag gaggtgggtg tccgcagcgc taacttcgcc gacaacgcca 2580
aggtgatcat cgtgcagctg aacgagagcg tggagatcaa ctgcaccgc gctagcgcgc 2640
actgcaacct gagccgcgcc aagtggaaag acaccctgaa caagatcgtg atcaagctgc 2700
gcgagcagtt cggcaacaag accatcgtgt tcaagcacag cagcggcggc gaccccgaga 2760
tcgtgacca cagcttcaac tgcggcgggc agttcttcta ctgcaacagc gctagcctgc 2820
cctgccgcat caagcagatc atcaacatgt ggcagaaggc gggcaaggcc atgtacgccc 2880
cccccatcag cggccagatc cgctgcagca gcaacatcac cggcctgctg ctgaccgcgc 2940
acggcggcaa cagcaacaac gagagcgaga tcttccgtcc gggcgggcgc gacatgcgcg 3000
acaactggcg cagcgagctg tacaagtaca aggtgggtgaa gatcgagccc ctgggcgtgg 3060
ccccaccaa ggccaagctt accgtccagg cccgccagct gctgagcggc atcgtgcagc 3120

```

```

agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag ctgaccgtgt 3180
ggggcatcaa gcagctgcag gcccgacccc tggcctgtga gcgctacctg aaggaccagc 3240
agcagctcga gcagatcttg aaccacacca cctggatgga gtgggaccgc gagatcaaca 3300
actacaccag cctgatccac agcctgatcg aggagagcca gaaccagcag gagaagaacg 3360
agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc aacatcacca 3420
actggctgtg gtacatcaag ctgttcatca tgatcgtggg cggcctgggt ggccctgcga 3480
tcgtgttcgc cgtgctgagc atctgaggat ccagatctgc tgtgccttct agttgccagc 3540
catctgttgt ttgccccctc cccgtgcctt ccttgaccct ggaagggtgc actcccactg 3600
tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc 3660
tgggggggtg ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg 3720
ctggggatgc ggtgggctct atgggtaccc aggtgctgaa gaattgacct ggttcctcct 3780
ggggcagaaa gaagcaggca catccccctt tctgtgacac accctgtcca cgccccctgt 3840
tcttagttcc agccccactc ataggacact catagctcag gagggtcccg ccttcaatcc 3900
caccgcctaa agtacttgga gcggtctctc cctccctcat cagcccacca aaccaaacct 3960
agcctccaag agtggaaga aattaaagca agataggcta ttaagtgcag agggagagaa 4020
aatgcctcca acatgtgagg aagtaatgag agaaatcata gaattttaag gccatgattt 4080
aaggccatca tggccttaat cttccgcttc ctcgctcact gactcgtgc gctcggtcgt 4140
tcggctgcgg cgagcgggat cagctcactc aaaggcggta atacggttat ccacagaatc 4200
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 4260
aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 4320
tcgacgctca agtcagagggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc 4380
ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 4440
cgcccttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 4500
ttcgggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccg 4560
ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaaagc acgacttatc 4620
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 4680
agagttcttg aagtgggtgg ctaactacgg ctacactaga agaacagtat ttggtatctg 4740
cgctctgctg aagccagtta ccttcggaaa aagagtgtgt agctcttgat ccggcaaaca 4800
aaccaccgct ggtagcgggt gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 4860
aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa 4920
ctcacgttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt 4980
aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 5040
ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat 5100
agttgcctga ctcggggggg gggggcgctg aggtctgcct cgtgaagaag gtgttgctga 5160
ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc cacggttgat 5220
gagagctttg ttgtagggtg accagtgtgt gatthtgaac ttttgctttg ccacggaacg 5280
gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag ttcgatttat 5340
tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct gccagtgtta caaccaatta 5400
accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt attcatatca 5460
ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga aaactcacgc 5520
aggcagttcc ataggatggc aagatcctgg tatcgggtct cgattccgac tcgtccaaca 5580
tcaatacaac ctattaattt cccctcgtca aaaataagggt tatcaagtga gaaatcacca 5640
tgagtgcga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt ccagacttgt 5700
tcaacaggcc agccattacg ctcgatcatc aaatcactcg catcaaccaa accgttattc 5760
attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg acaattacaa 5820
acaggaatcg aatgcaaccg gcgcaggaa actgccagcg catcaacaat attttccact 5880
gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc agtggtgagt 5940
aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg cataaattcc 6000
gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct acctttgcca 6060
tgtttcagaa acaactctgg cgcacggggt ttcccatata atcgatagat tgtcgcacct 6120
gattgcccga cattatcgcg agccatttta taccatata aatcagcatc catgttgga 6180
tttaatcgcg gcctcgagca agacgttttc cgttgaatat ggctcataac accccttgta 6240
ttactgttta tgtaagcaga cagttttatt gttcatgatg atatatTTTT atcttggtgca 6300
atgtaacatc agagattttg agacacaacg tggctttccc cccccccca ttattgaagc 6360
atttatcagg gttattgtct catgagcgga tacatatTTT aatgtattta gaaaaataaa 6420

```

caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 6480  
 attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tc 6532

<210> 70

<211> 7181

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 70

tccgcgcttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420  
 cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtac atctacgtat tagtcatcgc battaccatg gtgatgcggg tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttcccaaaat gtcgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900  
 agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatggggtc ccattttatta tttaaaatt cacatataca acaacgccgt ccccgctgcc 1380  
 cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500  
 agcggctcat ggtcgtcggc cagctccttg ctcttaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtg tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tccctttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga acaccatgcg 1920  
 cgtgaaggag aagtaccagc acctgtggcg ctggggctgg cgctggggca ccatgctgct 1980  
 gggcatgctg atgatctgca gcgccaccga gaagctgtgg gtgaccgtgt actacggcgt 2040  
 gcccgctgtg aaggaggcca ccaccaccct gctctgcgcc agcgacgcca aggcctacga 2100  
 caccgaggtg cacaacgtgt gggccaccca cgcctgcgtg cccaccgacc ccaaccccca 2160  
 ggaggtgggt ctggtgaacg tgaccgagaa cttcgacatg tggagaacg acatgggtgga 2220  
 gcagatgcac gaggacatca tcagcctgtg ggaccagagc ctgaagccct gcgtgaagct 2280  
 gacccccctg tgcgtgagcc tgaagtgcac cgacctgaag aacgacacca acaccaacag 2340  
 cagcagcggc cgcattgatc tggagaaggg cgagatcaag aactgcagct tcaacatcag 2400  
 caccagcatc cgcggcaagg tgcagaagga gtacgccttc ttctacaagc tggacatcat 2460



```

ccccatcgac aacgacacca ccagctacag cctgaccagc tgcaaacacca gcgtgatcac 2520
ccaggcctgc cccaaggtga gcttcgagcc catccccaac cactactgcg ccccgccgg 2580
cttcaccatc ctgaagtga aggacaagaa gttcaacggc aaggggccct gcaccaacgt 2640
gagcaccgtg cagtgcaccc acggcatccg ccccggtgtg agcaccagc tgctgctgaa 2700
cgaccgcctg gccgaggagg aggtggtgat ccgcagcgcc aacttcgccc acaacgcaa 2760
ggtgatcatc gtgcagctga acgagagcgt ggagatcaac tgcaccgcgc ccaacaacaa 2820
caccgcgaag agcatccaca tcggccccgg ccgcgccttc tacaccaccg gcgagatcat 2880
cggcgacatc cgccaggccc actgcaacct gagccgcgcc aagtggaaac acaccctgaa 2940
caagatcggt atcaagctgc gcgagcagtt cggcaacaag accatcggtg tcaagcacag 3000
cagcggcggc gaccccgaga tcgtgaccca cagcttcaac tgcggcggcg agttcttcta 3060
ctgcaacagc acccagctgt tcaacagcac ctggttcaac agcacctgga gcaccgaggg 3120
cagcaacaac accgagggca gcgacaccat caccctgccc tgccgcatca agcagatcat 3180
caacatgtgg cagaaggtgg gcaaggccat gtacgcccc cccatcagcg gccagatccg 3240
ctgcagcagc aacatcaccc gcctgctgct gacccgcgac ggcggcaaca gcaacaacga 3300
gagcgagatc ttccgcctgg gcggcggcga catgcgcgac aactggcgca gcgagctgta 3360
caagtacaag gtggtgaaga tcgagcccc gggcgtggcc cccaccaagg ccaagcgccg 3420
cgtggtgcag cgcgagaagc gcgccgtggg catcggcgcc ctgttcctgg gcttcctggg 3480
cgccgcggc agcaccatgg gcgccgccag catgacctg accgtgcagg cccgccagct 3540
gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca 3600
cctgctgcag ctgaccgtgt ggggcatcaa gcagctgcag gcccgcatcc tggcgtgga 3660
gcgctacctg aaggaccagc agctgctggg catctggggc tgcagcggca agcttatctg 3720
caccaccgcc gtgccctgga acgccagctg gagcaacaag agcctggagc agatctggaa 3780
ccacaccacc tggatggagt gggaccgcga gatcaacaac tacaccagcc tgatccacag 3840
cctgatcgag gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga 3900
caagtgggccc agcctgtgga actggttcaa catcaccaac tggctgtggt acatcaagct 3960
gttcatcatg atcgtgggcg gcctggtggg cctgcgcate gtgttcgccc tgctgagcat 4020
cgtgaaccgc gtgcgccagg gctacagccc cctgagcttc cagaccacc tgcccatccc 4080
ccgcggcccc gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg 4140
cacgatccgc ctggtgaacg gcagctgagg atccagatct gctgtgcctt ctagttgcca 4200
gccatctgtt gtttgccctt ccccgctgcc ttcttgacc ctggaagggt ccactcccac 4260
tgtcctttcc taataaaatg aggaaattgc atcgcatgtt ctgagtaggt gtcattctat 4320
tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 4380
tgctggggat gcggtgggct ctatgggtac ccagggtgct aagaattgac ccggttcctc 4440
ctgggccaga aagaagcagg cacatccctt tctctgtgac acaccctgtc cacgcccctg 4500
gttcttagtt ccagccccac tcataggaca ctcatagctc aggagggtc cgcttcaat 4560
cccaccgcct aaagtacttg gagcggtctc tccctccctc atcagcccac caaaccaaac 4620
ctagcctcca agagtgggaa gaaattaaag caagataggc tattaagtgc agaggagag 4680
aaaatgcctc caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatcat 4740
ggccttaate ttccgcttcc tcgctactg actcgctgcg ctcggtcggt cggtgcggc 4800
gagcgggtatc agtccactca aaggcggtaa tacggttatc cacagaatca ggggataacg 4860
caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt 4920
tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa 4980
gtcagagggt gcgaaaccgc acaggactat aaagatacca ggcgtttccc cctggaagct 5040
ccctcggtgc ctctcctgtt ccgacctgc cgcttaccgg atacctgtcc gcctttctcc 5100
cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg 5160
tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct 5220
tatccggtaa ctatcgtctt ggtccaacc cggtaagaca cgacttatcg ccactggcag 5280
cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga 5340
agtgggtggc taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga 5400
agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaaaaa accaccgctg 5460
gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag 5520
aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag 5580
ggattttggc catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat 5640
gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct 5700
taatcagtga ggcacctatc tcagcgatct gtctatttcc ttcattccata gttgcctgac 5760

```

```

tcggggggggg ggggcgctga ggtctgcctc gtgaagaagg tggtgctgac tcataccagg 5820
cctgaatcgc cccatcatcc agccagaaag tgagggagcc acggttgatg agagctttgt 5880
tgtaggtgga ccagttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt 5940
cgggaagatg cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc 6000
gccgtcccggt caagtcagcg taatgctctg ccagtggttac aaccaattaa ccaattctga 6060
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat 6120
accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 6180
taggatggca agatcctggt atcgggtctgc gattccgact cgtccaacat caatacaacc 6240
tattaatttc ccctcgtaaa aaataagggt atcaagttag aaatcaccat gagtgacgac 6300
tgaatccggt gagaatggca aaagcttatg catttctttc cagacttggt caacaggcca 6360
gccattacgc tcgtcatcaa aatcactcgc atcaacaaaa ccgttattca ttcgtgattg 6420
cgcctgagcg agacgaaata cgcgatcgtc gttaaaagga caattacaaa caggaatcga 6480
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata 6540
ttcttctaat acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcac 6600
atcaggagta cggataaaat gcttgatggt cggagaggc ataaattccg tcagccagtt 6660
tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa 6720
caactctggc gcacggggt tcccatataa tcgatagatt gtcgcacctg attgcccagc 6780
attatcgcg gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcg 6840
cctcgagcaa gacgtttccc gttgaatatg gctcataaca ccccttgat tactgtttat 6900
gtaagcagac agttttattg ttcattgatg tatattttta tcttggtgca tgtaacatca 6960
gagattttga gacacaacgt ggctttcccc ccccccccat tattgaagca tttatcaggg 7020
ttattgtctc atgagcggat acatatgtga atgtatttag aaaaataaac aaataggggt 7080
tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac 7140
attaacctat aaaaataggg gtatcacgag gccctttcgt c 7181

```

&lt;210&gt; 71

&lt;211&gt; 7493

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 71

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgccctgg tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatgt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140

```

taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	acaccatgcg	1920
cgtgaaggag	aagtaccagc	acctgtggcg	ctggggctgg	cgtgggggca	ccatgctgct	1980
gggcatgctg	atgatctgca	gcgccaccga	gaagctgtgg	gtgaccgtgt	actacggcgt	2040
ggccgtgtgg	aaggaggcca	ccaccaccct	gctctgcgcc	agcgacgcca	aggcctacga	2100
caccgaggtg	cacaacgtgt	gggccaccga	cgcctgcgtg	cccaccgacc	ccaaccccca	2160
ggaggtgggtg	ctggtgaacg	tgaccgagaa	cttcgacatg	tggagaagacg	acatgggtgga	2220
gcagatgcac	gaggacatca	tcagcctgtg	ggaccagagc	ctgaagccct	gcgtgaagct	2280
gacccccctg	tgcgtgagcc	tgaagtgcac	cgacctgaag	aacgacacca	acaccaacag	2340
cagcagcggc	cgcattgatca	tggagaaggg	cgagatcaag	aactgcagct	tcaacatcag	2400
caccagcatc	cgcggcaagg	tgcagaagga	gtacgccttc	ttctacaagc	tggacatcat	2460
ccccatcgac	aacgacacca	ccagctacag	cctgaccagc	tgcaacacca	gcgtgatcac	2520
ccaggcctgc	cccaagggtga	gcttcgagcc	catccccaac	cactactgcg	ccccgcgcg	2580
cttcaccatc	ctgaagtgc	aggacaagaa	gttcaacggc	aagggccctt	gcaccaacgt	2640
gagcaccgtg	cagtgcaccc	acggcatccg	ccccgtgggtg	agcaccagc	tgctgctgaa	2700
cgaccgcctg	gccgaggagg	aggtggtgat	cgcgagcgcc	aacttcgccc	acaacgcca	2760
ggtgatcatc	gtgcagctga	acgagagcgt	ggagatcaac	tgcacccgcc	ccaacaacaa	2820
caccgcgaag	agcatccaca	tgggccccgg	cgcgccttc	tacaccaccg	gcgagatcat	2880
cggcgacatc	cgccaggccc	actgcaacct	gagccgcgcc	aagtggaaacg	acaccctgaa	2940
caagatcgtg	atcaagctgc	gcgagcagtt	cggcaacaag	accatcgtgt	tcaagcacag	3000
cagcggcggc	gaccccgaga	tctgaccca	cagcttcaac	tgcggcggcg	agttcttcta	3060
ctgcaacagc	accagctgt	tcaacagcac	ctggttcaac	agcacctgga	gcaccgaggg	3120
cagcaacaac	accgagggca	gcgacaccat	caccctgccc	tgcgcgatca	agcagatcat	3180
caacatgtgg	cagaagggtg	gcaaggccat	gtacgcccc	cccatcagcg	gccagatccg	3240
ctgcagcagc	aacatcaccg	gcctgtgtgt	gacccgcgac	ggcggcaaca	gcaacaacga	3300
gagcgagatc	ttccgcctgg	gcggcggcga	catgcgcgac	aactggcgca	gcgagctgta	3360
caagtacaag	gtggtgaaga	tgcagccctt	gggcgtggcc	cccaccaagg	ccaagcgccg	3420
cgtggtgcag	cgcgagaagc	gcgccgtggg	catcggcgcc	ctgttcctgg	gcttcctggg	3480
cgcgcgcggc	agcaccatgg	gcgccgccag	catgaccctg	accgtgcagg	cccgccagct	3540
gctgagcggc	atcgtgcagc	agcagaacaa	cctgctgcgc	gccatcgagg	cccagcagca	3600
cctgctgcag	ctgaccgtgt	ggggcatcaa	gcagctgcag	gcccgcattc	tggccgtgga	3660
gcgctacctg	aaggaccagc	agctgctggg	catctggggc	tgcagcgga	agcttatctg	3720
caccaccgcc	gtgccctgga	acgccagctg	gagcaacaag	agcctggagc	agatctggaa	3780
ccacaccacc	tggatggagt	gggaccgcga	gatcaacaac	tacaccagcc	tgatccacag	3840
cctgatcgag	gagagccaga	accagcagga	gaagaacgag	caggagctgc	tggagctgga	3900
caagtggggc	agcctgtgga	actgggtcaa	catcaccaac	tggctgtggt	acatcaagct	3960
gttcatcatg	atcgtggggc	gcctgggtgg	cctgcgcatt	gtgttcgccc	tgctgagcat	4020
cgtgaaccgc	gtgcgccagg	gctacagccc	cctgagcttc	cagacccaac	tgcccatccc	4080
ccgcggcccc	gaccgccccg	agggcatcga	ggaggagggc	ggcgagcgcg	accgcgaccg	4140
cacgatccgc	ctggtgaacg	gcagcctggc	cctgatctgg	gacgacctgc	gcagcctgtg	4200
cctgttcagc	taccaccgcc	tgcgcgacct	gctgctgata	gtgaccgcga	tctgtggagct	4260
gctgggcccgc	cgcggctggg	aggccctgaa	gtactggtgg	aacctgctgc	agtactggag	4320
ccaggagctg	aagaacagcg	ccgtgagcct	gctgaacgcc	accgccatcg	ccgtggccga	4380
gggcaccgac	cgcgtgatcg	aggtggtgca	gggcgcctgc	cgcgccatcc	gccacatccc	4440



```

ccgccgcate cgccagggcc tggagcgcab cctgctgtga ggatccagat ctgctgtgcc 4500
ttctagttgc cagccatctg ttgtttgccc ctcccccgctg ccttccttga ccctggaagg 4560
tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcat gtctgagtag 4620
gtgtcattct attctggggg gtgggggtggg gcaggacagc aagggggagg attggaaga 4680
caatagcagg catgctgggg atgcggtggg ctctatgggt acccaggtgc tgaagaattg 4740
acccggttcc tcctgggcca gaaagaagca ggcacatccc cttctctgtg acacaccctg 4800
tccacgcccc tggttcttag ttccagcccc actcatagga cactcatagc tcaggagggc 4860
tccgccttca atcccacccg ctaaagtact tggagcggtc tctccctccc tcatcagccc 4920
accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt 4980
gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt 5040
taaggccatc atggccttaa tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 5100
ttcggtgctg gcgagcggta tcagctcact caaaggcggt aatacgggtta tccacagaat 5160
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 5220
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 5280
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 5340
cccctggaag ctccctcgct cgctctcctg ttccgacctt gccgcttacc ggatacctgt 5400
ccgcctttct cccttcggga agcgtggcgc tttctcatag ctacgctgt aggtatctca 5460
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc 5520
accgctgctc cttatccggt aactatcgtc ttgagtccta cccggtaga cagacttat 5580
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 5640
cagagttctt gaagtgggtg cctaactacg gctacactag aagaacagta tttggtatct 5700
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 5760
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 5820
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 5880
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 5940
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggctctgaca 6000
gttaccaatg cttaatcagt gaggcacct tctcagcgat ctgtctatct cgttcatcca 6060
tagttgcctg actcgggggg ggggggcgct gaggtctgcc tcgtgaagaa ggtgttgctg 6120
actcatacca ggcctgaatc gccccatcat ccagccagaa agtgaggag ccacggttga 6180
tgagagcttt gttgtagggt gaccagttgg tgattttgaa cttttgcttt gccacggaac 6240
ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctacgcaaaa gttcgattta 6300
ttcaacaaag ccgcgcgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt 6360
aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc 6420
aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc 6480
gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga ctctcccaac 6540
atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg agaaatcacc 6600
atgagtgcag actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg 6660
ttcaacaggc cagccattac gctcgtcac aaaatcactc gcatcaacca aaccgttatt 6720
catctgctgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaaag gacaattaca 6780
aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa tattttcacc 6840
tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg cagtgggtgag 6900
taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaattc 6960
cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc 7020
atgtttcaga aacaactctg gcgcateggg ctcccatatc aatcgataga ttgtcgcacc 7080
tgattgcccc acattatcgc gagcccatth ataccatata aaatcagcat ccatgttgga 7140
atttaatcgc ggcctcgagc aagacgtttc ccgttgataa tggctcataa cacccttgt 7200
attactgttt atgtaagcag acagttttat tgttcatgat gatataatth tatcttgtgc 7260
aatgtaacat cagagatttt gagacacaac gtggctttcc ccccccccc attattgaag 7320
catttatcag ggttatgttc tcatgagcgg atacatattt gaatgtattt agaaaaataa 7380
acaaataggg gttccgcgca catthccccg aaaagtgcc cctgacgtct aagaaaccat 7440
tattatcatg acattaacct ataaaaatag gcgtatcacg aggccttttc gtc 7493

```

&lt;211&gt; 7186

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 72

```
tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccggtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggatgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaagggtgag cttcgagccc atccccaacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggcccttg caccaacgtg 2640
agcaccgtgc agtgacacca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820
```

```

acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880
atcggcaaca tgcgccagge ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcggt acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tggcctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaagggtgt gaagatcgag cccctgggcg tggccccac caaggccaag 3420
cgccgcgtgg tgcagcgcgga gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480
ctgggcgccc cgggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccg 3540
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgata 3840
cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960
aagctgttca tcatgatcgt gggcgccctg gtgggcctgc gcacgtgtt cgcctgctg 4020
agcatcgtga accgcgtgcg ccagggtac agccccctga gcttccagac ccacctgcc 4080
atcccccgcg gccccgaccg ccccgagggc atcgaggagg agggcgggcg gcgcgaccgc 4140
gaccgcacga tccgcctggt gaacggcagc tgaggatcca gatctgctgt gccttctagt 4200
tgccagccat ctgttgtttg cccctcccc gtgccttctt tgaccctgga aggtgccact 4260
cccactgtcc ttctctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 4320
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggatggga agacaatagc 4380
aggcatgctg gggatgcggt gggctctatg ggtaccaggg tgctgaagaa ttgaccgggt 4440
tcctcctggg ccagaaagaa gcaggcacat ccccttctct gtgacacacc ctgtccacgc 4500
ccctggttct tagttccagc cccactcata ggacactcat agctcaggag ggctccgcct 4560
tcaatccac ccgctaaagt acttgagcgc gtctctccct cctcatcag cccaccaaac 4620
caaacctagc ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg 4680
gagagaaaat gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc 4740
atcatggcct taatcttccg ctctctcgt cactgactcg ctgcgctcgg tcgttcgggt 4800
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4860
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4920
cgcgctgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 4980
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 5040
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 5100
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 5160
gtaggtcgtt cgctccaage tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg 5220
cgcttatcc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact 5280
ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcggtg ctacagagtt 5340
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct 5400
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 5460
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 5520
tcaagaagat cctttgatct ttctacggg gtctgacgct cagtggaaacg aaaactcacg 5580
ttaagggatt ttggatcatg gattatcaaa aaggatcttc acctagatcc ttttaaatta 5640
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 5700
atgcttaate agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5760
ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaagggtgtg ctgactcata 5820
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5880
tttggtgtag gtggaccagt tgggtgatttt gaacttttgc tttgccacgg aacggtctgc 5940
gttgctcgga agatgcgtga tctgacctt caactcagca aaagtctgat ttattcaaca 6000
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 6060
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 6120

```



```

tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgagggcag 6180
ttccatagga tggcaagatc ctggtatcgg tctgcgatcc cgactcgtcc aacatcaata 6240
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 6300
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 6360
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 6420
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 6480
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 6540
ggatattctt ctaataacctg gaatgctgtt ttcccgggga tcgcagtggg gagtaacct 6600
gcatcatcag gagtacggat aaaatgcttg atggtcgga gaggcataaa ttccgtcagc 6660
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6720
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6780
ccgacattat cgcgagccca ttataacca tataaatcag catccatgtt ggaatttaat 6840
cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccctt tgtattactg 6900
tttatgtaag cagacagttt battgttcat gatgatatat ttttatcttg tgcaatgtaa 6960
catcagagat ttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 7020
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 7080
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 7140
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 7186

```

&lt;210&gt; 73

&lt;211&gt; 7003

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 73

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgcctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatatata acaacgcctg cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggtc ccatgcctcc 1500

```

agcggtcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgccaccgag	aagctgtggg	tgaccgtgta	ctacggcgtg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcaa	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgacce	caacccccag	2160
gaggtggtgc	tgggtgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcgcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaagg	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtgag	cttcgagccc	atccccaacc	actactgcgc	ccccgccggc	2580
ttcaccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	ccggcccctg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggcagcctgg	ccgaggagga	ggtggtgata	cgcagcgtga	acttcaccga	caacgccaag	2760
accatcatcg	tgcagctgaa	caccagcgtg	gagatcaact	gcaccgccc	caacaacaac	2820
acccgcaagc	gaattcgcat	ccagcgcggc	cccgcccgcg	ccttcgtgac	catcggaag	2880
atcggcaaca	tgcgccaggc	ccactgcaac	atcagccgcg	ccaagtggaa	caacaccctg	2940
aagcagatcg	acagcaagct	gcgcgagcag	ttcggcaaca	acaagaccat	catcttcaag	3000
cagagcagcg	gcggcgacce	cgagatcgtg	acccacagct	tcaactgcgg	cggcgagttc	3060
ttctactgca	acagcaccca	gctgttcaac	agcacctgg	tcaacagcac	ctggagcacc	3120
gagggcagca	acaacaccga	gggcagcgac	accatcacc	tgccctgccg	catcaagcag	3180
atcatcaaca	tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	3240
atccgctgca	gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	3300
aacgagagcg	agatcttccg	cctggggcgg	ggcgacatgc	gcgacaactg	gcgcagcgag	3360
ctgtacaagt	acaaggtgg	gaagatcgag	cccctggcg	tggccccac	caaggccaag	3420
cttaccgtgc	aggcccgcga	gctgctgagc	ggcatcgtgc	agcagcagaa	caacctgctg	3480
cgcgccatcg	aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	3540
caggcccgcga	tcctggccgt	ggagcgctac	ctgaaggacc	agcagctgga	gcagatctgg	3600
aaccacacca	cctggatgga	gtgggaccgc	gagatcaaca	actacaccag	cctgatccac	3660
agcctgatcg	aggagagcca	gaaccagcag	gagaagaacg	agcaggagct	gctggagctg	3720
gacaagtggg	ccagcctgtg	gaactggttc	aacatcacca	actggctgtg	gtacatcaag	3780
ctgttcatca	tgatcgtggg	cggcctgggt	ggcctgcgca	tcgtgttcgc	cgtgctgagc	3840
atcgtgaacc	gcgtgcgcca	gggtacagc	cccctgagct	tcagaccca	cctgcccac	3900
ccccgcggcc	ccgaccgccc	cgagggcatc	gaggaggagg	gcggcgagcg	cgaccgcgac	3960
cgcacgatcc	gcctgggtgaa	cggcagctga	ggatccagat	ctgctgtgcc	ttctagtgtc	4020
cagccatctg	ttgtttgccc	ctccccctg	ccttccttga	ccctggaagg	tgccactccc	4080
actgtccttt	cctaataaaa	tgaggaaatt	gcacgcatt	gtctgagtag	gtgtcattct	4140
attctggggg	gtggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	4200
catgctgggg	atgcggtggg	ctctatgggt	accaggtgc	tgaagaattg	acccggttcc	4260
tcctgggcca	gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	tcacgcccc	4320
tggttcttag	ttccagcccc	actcatagga	cactcatagc	tcaggagggc	tcgccttca	4380
atcccacccg	ctaaagtact	tggagcggtc	tctccctccc	tcacagccc	accaaaccac	4440
acctagcctc	caagagtggg	aagaaattaa	agcaagatag	gctattaagt	gcagagggag	4500
agaaaatgcc	tccaacatgt	gaggaagtaa	tgagagaaat	catagaattt	taaggccatc	4560
atggccttaa	tcttccgctt	cctcgtcac	tgactcgtg	cgtcgggtcg	ttcggctgcg	4620
gcgagcggtg	tcagctcact	caaaggcgg	aatacggtta	tccacagaat	caggggataa	4680
cgcaggaaaag	aacatgtgag	caaaaaggcca	gcaaaaaggcc	aggaaccgta	aaaaggccgc	4740
gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	4800

```

aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag 4860
ctccctcgtg cgctctcctg ttccgacctt gccgcttacc ggatacctgt ccgcctttct 4920
cccttcggga agcgtggcgc ttctcatag ctacgctgt aggtatctca gttcgggtga 4980
ggtcggttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc accgctgcgc 5040
cttatccggt aactatcgtc ttgagtccaa cccggtaaga cagacttat cgccactggc 5100
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt 5160
gaagtgggtg cctaactacg gctacactag aagaacagta ttggtatct gcgctctgct 5220
gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc 5280
tggtagcggg ggtttttttt ttgcaagca gcagattacg cgcagaaaaa aaggatctca 5340
agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaaacgaaa actcacgtta 5400
agggattttg gtcattgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa 5460
atgaagtttt aaatcaatct aaagtatata tgagtaaact tggctctgaca gttaccaatg 5520
cttaatcagt gaggcacctt tctcagcgat ctgtctatct cgttcattcca tagttgcctg 5580
actcgggggg gggggggcgt gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca 5640
ggcctgaatc gccccatcat ccagccagaa agtgaggag ccacggttga tgagagcttt 5700
gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac ggtctgcgtt 5760
gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa gttcgattta ttcaacaaag 5820
ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt aaccaattct 5880
gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc aggattatca 5940
ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc gaggcagttc 6000
cataggatgg caagatcctg gtatcgggtc gcgattccga ctcgtccaac atcaatacaa 6060
cctattaatt tcccctcgtc aaaaataagg ttatcaagtg agaaatcacc atgagtgcgc 6120
actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg ttcaacaggc 6180
cagccattac gctcgtcatc aaaatcactc gcatcaacca aaccggttatt cattcgtgat 6240
tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaa gacaattaca aacaggaatc 6300
gaatgcaacc ggcgaggaa cactgccagc gcatcaacaa tattttcacc tgaatcagga 6360
tattcttcta atacctggaa tgctgttttc ccggggatcg cagtgggtgag taaccatgca 6420
tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaattc cgtcagccag 6480
tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc atgtttcaga 6540
aacaactctg gcgcacggg cttcccatac aatcgataga ttgtcgcacc tgattgcccg 6600
acattatcgc gagcccatct ataccatata aaatcagcat ccatgttgga atttaatcgc 6660
ggcctcgagc aagacgtttc ccgttgaata tggctcataa cacccttgtt attactgttt 6720
atgtaagcag acagttttat tgttcattgat gatataatct tatcttgtgc aatgtaacat 6780
cagagatttt gagacacaac gtggctttcc ccccccccc attattgaag catttatcag 6840
ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 6900
gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaacatc tattatcatg 6960
acattaacct ataaaaatag gcgtatcacg aggccttttc gtc 7003

```

&lt;210&gt; 74

&lt;211&gt; 7186

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 74

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360

```

```

ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattgg 420
cccgctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggtctg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggte ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggtcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctcttct 1860
tgcaatcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga cccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgtgtctg 1980
ggcatgctga tgatctgcag cggcaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcca ggccctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtgggtg tggtcgacgt gaccgagaac ttcgacatgt ggaagaacga catgggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacacca caaccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaagg actgtagctt cgacatcagc 2400
accagcatcc gcggcaagg gcaagaaggg tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcgaaactag tgtgatcacc 2520
caggcctgcc ccaagggtgag ctctcagacc atccccacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa cgacaagacc ttcgacggca ccgggcccctg caccgacgtg 2640
agcaccgtgc agtgacacca cggcatccgc cccgtgggtg gcacccagct gctgtctgac 2700
ggcagcctgg ccgaggagga ggtgggtgat cgcagcgtcg acttcaccga caacgccaag 2760
accatcatcg tgcagggtga caccacggtg gagatcgatt gcacccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880
atcggcaaca tgcgccaggc cactgcgat atcagccgcg ccaagtggga caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaacg acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgcg acagacacca gctgttcaac agcacctggg tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcacc tgcctgccc catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcgatat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaagggtgg gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
cgccgcgtgg tgcagcgcga gaagcgcgcc gtgggcatcg gcgcctgtt cctgggcttc 3480
ctgggcgcgg ccggcagcac catgggcgcc gccagcatga cctgacccgt gcaggcccgc 3540
cagctgctga gcggcatcgt gcagcagcag acaaacctgc tgcgcgccat cgaggcccag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggccc catcctggcc 3660

```



gtggagcgct	acctgaagga	ccagcagctg	ctgggcatct	ggggctgcag	cggaagctt	3720
atctgcacca	ccgccgtgcc	ctggaacgcc	agctggagca	acaagagcct	ggagcagatc	3780
tggaaaccaca	ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgatc	3840
cacagcctga	tcgaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	3900
ctggacaagt	gggccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtacatc	3960
aagctgttca	tcatgatcgt	gggcggcctg	gtgggcctgc	gcatcgtgtt	cgccgtgctg	4020
agcatcgtga	accgctgctg	ccagggctac	agccccctga	gcttccagac	ccacctgccc	4080
atcccccgcg	gccccgaccg	ccccgagggc	atcgaggagg	agggcgggcg	gcgcgaccgc	4140
gaccgcacga	tccgcctggg	gaacggcagc	tgaggatcca	gatctgctgt	gccttctagt	4200
tgccagccat	ctgttggttg	cccccccccc	gtgccttcct	tgacctgga	aggtgccact	4260
cccactgtcc	tttcctaata	aaatgaggaa	attgcacgc	attgtctgag	taggtgtcat	4320
tctattctgg	ggggtggggg	ggggcaggac	agcaaggggg	aggattggga	agacaatagc	4380
aggcatgctg	gggatgcggg	gggctctatg	ggtaccagag	tgctgaagaa	ttgacctggg	4440
tcctcctggg	ccagaaagaa	gcaggcacat	ccccctctct	gtgacacacc	ctgtccacgc	4500
ccctgggtct	tagttccagc	cccactcata	ggacactcat	agctcaggag	ggctccgcct	4560
tcaatcccac	ccgctaaagt	acttggagcg	gtctctccct	ccctcatcag	cccaccaaac	4620
caaacctagc	ctccaagagt	gggaagaaat	taaagcaaga	taggctatta	agtgcagagg	4680
gagagaaaat	gcctccaaca	tgtgaggaag	taatgagaga	aatcatagaa	ttttaaggcc	4740
atcatggcct	taatcttccg	cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	4800
gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	4860
taacgcagga	aagaacatgt	gagcaaaaagg	ccagcaaaaag	gccaggaacc	gtaaaaaggc	4920
cgcggtgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	4980
ctcaagtcat	aggtggcgaa	accgcagagg	actataaaga	taccaggcgt	ttccccctgg	5040
aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	5100
tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggatct	tcagttcggg	5160
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	5220
cgcttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	5280
ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	5340
cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtatttggtg	tctgcgctct	5400
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	5460
cgctggtagc	gggtggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	5520
tcaagaagat	cctttgatct	ttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	5580
ttaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaaatta	5640
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttgggtctg	acagttacca	5700
atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	5760
ctgactcggg	gggggggggg	gctgaggtct	gcctcgtgaa	gaagggtgtg	ctgactcata	5820
ccaggcctga	atcgccccat	catccagcca	gaaagtgagg	gagccacggg	tgatgagagc	5880
tttggtgtag	gtggaccagt	tgggtgatttt	gaacttttgc	tttgccacgg	aacgggtctgc	5940
gttgctcggg	agatgcgtga	tctgatcctt	caactcagca	aaagttcgat	ttattcaaca	6000
aagccgcggt	cccgtcaagt	cagcgtaatg	ctctgccagt	gttacaacca	attaaccaat	6060
tctgattaga	aaaactcatc	gagcatcaaa	tgaaactgca	atttatctcat	atcaggatta	6120
tcaataccat	atttttgaaa	aagccgtttc	tgtaatgaag	gagaaaactc	accgaggcag	6180
ttccatagga	tggcaagatc	ctgggtatcg	tctgcgattc	cgactcgtcc	aacatcaata	6240
caacctatta	atttccctc	gtcaaaaata	aggttatcaa	gtgagaaatc	accatgagtg	6300
acgactgaat	ccggtgagaa	tggcaaaaagc	ttatgcattt	ctttccagac	ttgttcaaca	6360
ggccagccat	tacgctcgtc	atcaaaaatca	ctcgcatcaa	ccaaaccggt	attcatctgt	6420
gattgcgctt	gagcgagacg	aaatacgcga	tcgctgttaa	aaggacaatt	acaaacagga	6480
atcgaatgca	accggcgagc	gaacactgcc	agcgcatcaa	caatattttc	acctgaatca	6540
ggatatctct	ctaatacctg	gaatgctgtt	ttcccgggga	tcgcagtggg	gagtaaccat	6600
gcatcatcag	gagtacggat	aaaatgcttg	atgggtcgga	gaggcataaa	ttccgtcagc	6660
cagtttagtc	tgaccatctc	atctgtaaca	tcattggcaa	cgctaccttt	gccatgtttc	6720
agaaacaact	ctggcgcac	gggtttccca	tacaatcgat	agattgtcgc	acctgattgc	6780
ccgacattat	cgcgagccca	tttatacca	tataaatcag	catccatggt	ggaatttaat	6840
cgcggcctcg	agcaagacgt	ttcccgttga	atatggctca	taacacccct	tgtattactg	6900
tttatgtaag	cagacagtct	tattgttcat	gatgatatat	ttttatcttg	tgcaatgtaa	6960

```

catcagagat tttgagacac aacgtggctt tccccccccc cccattattg aagcatttat 7020
caggggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 7080
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 7140
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 7186

```

&lt;210&gt; 75

&lt;211&gt; 7186

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 75

```

tcgcgcggtt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatc tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcggttac gtgttcogga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcggtg 2040
ccgctgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcca ggccctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggtcgacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340

```



```

agcagcggcc gcatgatcat ggagaagggc gagatcaagg actgtagctt cgacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcgaaactag tgtgatcacc 2520
caggcctgcc ccaaggtgag cttcgagccc atccccaacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa cgacaagacc ttcgacggca ccggggccctg caccgacgtg 2640
agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtcg acttcaccga caacgccaag 2760
accatcatcg tgcaggtgga caccacggtg gagatcgatt gcacccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggaag 2880
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttccggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcggt acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggcccccac caaggccaag 3420
cgccgcgtgg tgcagcgcga gaagcgcgcc gtgggcatcg gcgccctgtt cctgggcttc 3480
ctgggcggcc ccggcagcac catgggcgcc gccagcatga ccctgaccgt gcaggcccg 3540
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840
cacagcctga tcgaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960
aagctgttca tcatgatcgt gggcggcctg gtgggcctgc gcacgtgtt cgcctgctg 4020
agcatcgtga accgcgtgcg ccagggtctac agccccctga gcttccagac ccacctgccc 4080
atcccccgcg gccccgaccg ccccgagggc atcgaggagg agggcggcga gcgcgaccgc 4140
gaccgcacga tccgcctggt gaacggcagc tgaggatcca gatctgetgt gccttctagt 4200
tgccagccat ctgttgtttg cccctcccc gtgccttctc tgaccctgga aggtgccact 4260
cccactgtcc ttctctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 4320
tctattctgg ggggtggggg ggggcaggac agcaaggggg aggtattgga agacaatagc 4380
aggcatgctg gggatgcggt gggctctatg ggtaccagc tgctgaagaa ttgaccgggt 4440
tctcctggg ccagaaagaa gcaggcacat ccccttctct gtgacacacc ctgtccacgc 4500
ccctgggtct tagttccagc cccactcata ggacactcat agctcaggag ggctccgcct 4560
tcaatccac ccgctaaagt acttgagcg gtctctccct cctcatcag cccaccaaac 4620
caaacctagc ctccaagagt gggaagaaat taaagcaaga taggctatta agtgcagagg 4680
gagagaaaat gcctccaaca tgtgaggaag taatgagaga aatcatagaa ttttaaggcc 4740
atcatggcct taatcttccg ctccctcgct cactgactcg ctgcgctcgg tcgttcgggt 4800
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4860
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4920
cgcgctgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg 4980
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 5040
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 5100
tctcccttcg ggaagcgtgg cgttttctca tagctcacgc tgtaggtatc tcagttcgggt 5160
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 5220
cgcttatcc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact 5280
ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcgggt ctacagagtt 5340
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggt tctgcgctct 5400
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 5460
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 5520
tcaagaagat cctttgatct ttctacggg gtctgacgct cagtggaacg aaaactcacg 5580
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatba 5640

```

```

aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttgggtctg acagttacca 5700
atgcttaate agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5760
ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 5820
ccaggcctga atcgcccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5880
tttgtttag gtggaccagt tgggtatttt gaacttttgc tttgccacgg aacggtctgc 5940
gttgtcggga agatgcgtga tctgacccct caactcagca aaagtctgat ttattcaaca 6000
aagccgcccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 6060
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 6120
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 6180
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 6240
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 6300
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 6360
ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccggt attcattcgt 6420
gattgcgcct gagcgagacg aaatacgaga tgcgtgttaa aaggacaatt acaaacagga 6480
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatatcttc acctgaatca 6540
ggatattctt ctaatacctg gaatgctgtt ttcccgggga tgcagtggt gagtaaccat 6600
gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6660
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgttct 6720
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6780
ccgacattat cgcgagccca ttatatacca tataaatcag catccatgtt ggaatttaac 6840
cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccctt tgtattactg 6900
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6960
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 7020
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 7080
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 7140
atgacattaa cctataaaaa taggcgtatc acgagggcct ttcgtc 7186

```

&lt;210&gt; 76

&lt;211&gt; 7498

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 76

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020

```

tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	cottatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttacaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagtttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcgagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatacgcg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgcaccgag	aagctgtggg	tgaccgtgta	ctacggcgctg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcaa	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtgggtgc	tgggtgaacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaaga	actgcagctt	caacatcagc	2400
accagcatcc	gcggcaagggt	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcaacaccag	cgtgatcacc	2520
caggcctgcc	ccaaggtgag	cttcgagccc	atccccaacc	actactgcgc	ccccgccggc	2580
ttcaccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	ccggccccctg	caccaacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccgtggtga	gcaccagct	gctgctgaac	2700
ggcagcctgg	ccgaggagga	ggtgggtgatc	cgcagcgtga	acttcaccga	caacgccaag	2760
accatcatcg	tgcagctgaa	caccagcgtg	gagatcaact	gcaccgccc	caacaacaac	2820
acccgcaagc	gaattcgcat	ccagcgcggc	cccggccgcg	ccttcgtgac	catcggcaag	2880
atcggcaaca	tgcgccaggc	ccactgcaac	atcagccgcg	ccaagtggaa	caacaccctg	2940
aagcagatcg	acagcaagct	gcgcgagcag	ttcggcaaca	acaagaccat	catcttcaag	3000
cagagcagcg	gcggcgaccc	cgagatcgtg	accacagct	tcaactgcgg	cggcgagttc	3060
ttctactgca	acagcaccca	gctgttcaac	agcacctggt	tcaacagcac	ctggagcacc	3120
gagggcagca	acaacaccga	gggcagcgac	accatcacc	tgccctgccc	catcaagcag	3180
atcatcaaca	tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	3240
atccgctgca	gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	3300
aacgagagcg	agatcttccg	cctgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3360
ctgtacaagt	acaagggtgt	gaagatcgag	cccctgggcg	tggccccac	caaggccaag	3420
cgcgcgctgg	tgcagcgcga	gaagcgcgcc	gtgggcatcg	gcgcctgtt	cctgggcttc	3480
ctgggcgccc	ccggcagcac	catgggcgcc	gccagcatga	ccctgaccgt	gcaggccccgc	3540
cagctgctga	gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggcccag	3600
cagcacctgc	tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccc	catcctggcc	3660
gtggagcgct	acctgaagga	ccagcagctg	ctgggcatct	ggggctgcag	cggcaagctt	3720
atctgcacca	ccgccgtgcc	ctggaacgcc	agctggagca	acaagagcct	ggagcagatc	3780
tggaaaccaca	ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgatc	3840
cacagcctga	tcgaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	3900
ctggacaagt	gggccagcct	gtggaactgg	ttcaacatca	ccaactggct	gtggtacatc	3960
aagctgttca	tcatgatcgt	gggcggcctg	gtgggcctgc	gcatacgtgt	cgcctgctg	4020
agcatcgtga	accgcgtgcg	ccagggtctac	agccccctga	gcttccagac	ccacctgccc	4080
atcccccgcg	gccccgaccg	ccccgagggc	atcgaggagg	agggcggcga	gcgcgaccgc	4140
gaccgcacga	tccgcctggt	gaacggcagc	ctggccctga	tctgggacga	cctgcgcagc	4200
ctgtgcctgt	tcagctacca	ccgcctgcgc	gacctgctgc	tgatcgtgac	ccgcatcgtg	4260
gagctgctgg	gccgcgcg	ctgggaggcc	ctgaagtact	ggtggaacct	gctgcagtac	4320



tggagccagg	agctgaagaa	cagcgccgtg	agcctgctga	acgccaccgc	catcgccgtg	4380
gccgagggca	ccgaccgcgt	gacgaggtg	gtgcagggcg	cctgccgcgc	catccgccac	4440
atcccccgcc	gcatccgcca	gggctggag	cgcacctgc	tgtgaggatc	cagatctgct	4500
gtgccttcta	gttgccagcc	atctgttgtt	tgccctccc	ccgtgccttc	cttgaccctg	4560
gaaggtgcca	ctcccactgt	cctttcctaa	taaaatgagg	aaattgcatc	gcattgtctg	4620
agtaggtgtc	attctattct	ggggggtggg	gtggggcagg	acagcaaggg	ggaggattgg	4680
gaagacaata	gcaggcatgc	tggggatgcg	gtgggctcta	tgggtaccca	ggtgctgaag	4740
aattgacccg	gttcctcctg	ggccagaaa	aagcaggcac	atccccctct	ctgtgacaca	4800
ccctgtccac	gcccctgggt	cttagttcca	gccccactca	taggacactc	atagctcagg	4860
agggctccgc	cttcaatccc	acccgctaaa	gtacttggag	cggctctctc	ctccctcatc	4920
agcccaccaa	accaaacct	gcctccaaga	gtgggaagaa	attaaagcaa	gataggctat	4980
taagtgcaga	gggagagaaa	atgcctccaa	catgtgagga	agtaatgaga	gaaatcatag	5040
aattttaagg	ccatcatggc	cttaatcttc	cgcttccctc	ctcactgact	cgctgcgctc	5100
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatac	ggttatccac	5160
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	5220
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgccccctg	acgagcatca	5280
caaaaatcga	cgctcaagtc	agagggtggc	aaacccgaca	ggactataaa	gataccaggc	5340
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	5400
cctgtccgcc	tttctccctt	cgggaaagcg	ggcgctttct	catagctcac	gctgtaggta	5460
tctcagttcg	gtgtaggtcg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	5520
gcccgaccgc	tgcgccttat	ccggtaaact	tcgtcttgag	tccaacccgg	taagacacga	5580
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	5640
tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	5700
tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	5760
caaacaacc	accgctggta	gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	5820
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	5880
cgaaaactca	cgttaaggga	ttttgggtcat	gagattatca	aaaaggatct	tcacctagat	5940
ccttttaaat	taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	6000
tgacagttac	caatgcttaa	tcagtgaggc	acctatctca	gcgatctgtc	tatttcgttc	6060
atccatagtt	gcctgactcg	gggggggggg	gcgctgaggt	ctgcctcgty	aagaaggtyt	6120
tgctgactca	taccaggcct	gaatcgcccc	atcatccagc	cagaaagtga	gggagccacg	6180
gttgatgaga	gctttgttgt	aggtggacca	gttggtgatt	ttgaactttt	gctttgccac	6240
ggaacggtct	gcgttgctcg	gaagatgcgt	gatctgatcc	ttcaactcag	caaaagttcg	6300
atttattcaa	caaagccgcc	gtcccgtcaa	gtcagcgtaa	tgctctgcca	gtgttacaac	6360
caattaacca	attctgatta	gaaaaactca	tcgagcatca	aatgaaactg	caatttatte	6420
atatcaggat	tatcaatacc	atatttttga	aaaagccgtt	tctgtaatga	aggagaaaac	6480
tcaccgaggg	agttccatag	gatggcaaga	tcctgggtatc	ggctctcgat	tccgactcgt	6540
ccaacatcaa	tacaacctat	taatttcccc	tcgtcaaaaa	taaggttatc	aagtgagaaa	6600
tcaccatgag	tgacgactga	atccggtgag	aatggcaaaa	gcttatgcat	ttctttccag	6660
acttgttcaa	caggccagcc	attacgctcg	tcatcaaaat	cactcgcac	aaccaaaccg	6720
ttattcatte	gtgattgcgc	ctgagcgaga	cgaaatacgc	gatcgctggt	aaaaggacaa	6780
ttacaaacag	gaatcgaatg	caaccggcgc	aggaacactg	ccagcgcac	aacaatattt	6840
tcacctgaat	caggatatte	ttctaatacc	tggaaatgctg	ttttcccggg	gatcgcagtg	6900
gtgagtaacc	atgcatcatc	aggagtacgg	ataaaatgct	tgatgggtcgg	aagaggcata	6960
aattccgtca	gccagtttag	tctgaccatc	tcactctgtaa	catcattggc	aacgctacct	7020
ttgccatggt	tcagaaacaa	ctctggcgca	tcgggcttcc	catacaatcg	atagattgtc	7080
gcacctgatt	gcccgcacatt	atcgcgagcc	catttatacc	catataaatc	agcatccatg	7140
ttggaattta	atcgcggcct	cgagcaagac	gtttcccgtt	gaatatggct	cataacaccc	7200
cttgatttac	tgtttatgta	agcagacagt	tttattgttc	atgatgatat	atttttatct	7260
tgtgcaatgt	aacatcagag	atthttgagac	acaacgtggc	tttccccccc	ccccatttat	7320
tgaagcattt	atcagggtta	ttgtctcatg	agcggataca	tatttgaatg	tatttagaaa	7380
aataaaca	taggggttcc	gcgcacattt	ccccgaaaag	tgccacctga	cgtctaagaa	7440
accattatta	tcatgacatt	aacctataaa	aataggcgta	tcacgaggcc	ctttcgtc	7498

&lt;210&gt; 77

&lt;211&gt; 7315

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 77

```

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcggt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc lagatcgccct gagacgccat ccacgctggt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgcggt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagatc tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgcgtccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat ggggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccggtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtgggtg tgggtgaacgt gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaaga actgcagctt caacatcagc 2400
accagcatcc gcggcaaggt gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcaacaccag cgtgatcacc 2520
caggcctgcc ccaaggtgag cttcgagccc atccccaaac actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa caacaagacc ttcaacggca ccggccctg caccaacgtg 2640
agcaccgtgc agtgacacca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700

```

```

ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtga acttcaccga caacgccaag 2760
accatcatcg tgcagctgaa caccagcgtg gagatcaact gcacccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120
gagggcagca acaacaccga gggcagcgac accatcacc tgccttgcgg catcaagcag 3180
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300
aacgagagcg agatcttcg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360
ctgtacaagt acaaggtggt gaagatcgag cccctgggcg tggccccac caaggccaag 3420
cttaccgtgc aggcccgcca gctgctgagc ggcacgtgc agcagcagaa caacctgctg 3480
cgcgccatcg aggcccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 3540
caggcccgca tcctggccgt ggagcgtac ctgaaggacc agcagctgga gcagatctgg 3600
aaccacacca cctggatgga gtgggaccgc gagatcaaca actacaccag cctgatccac 3660
agcctgatcg aggagagcca gaaccagcag gagaagaacg agcaggagct gctggagctg 3720
gacaagtggg ccagcctgtg gaactggttc aacatcacca actggctgtg gtacatcaag 3780
ctgttcatca tgatcgtggg cggcctggtg ggcctgcgca tcgtgttcgc cgtgctgagc 3840
atcgtgaacc gcgtgcgcca gggctacagc cccctgagct tccagaccca cctgcccac 3900
ccccgcggcc ccgaccgccc cgagggcatc gaggaggagg gcggcgagcg cgaccgcgac 3960
cgcacgatcc gcctggtgaa cggcagcctg gccctgatct gggacgacct gcgcagcctg 4020
tgctgttca gctaccaccg cctgcgcgac ctgctgctga tcgtgacctg catcgtggag 4080
ctgctgggccc gccgcggctg ggaggccctg aagtactggt ggaacctgct gcagtactgg 4140
agccaggagc tgaagaacag cgccgtgagc ctgctgaacg ccaccgccat cgccgtggcc 4200
gagggcaccg accgcgtgat cgaggtggtg cagggcgccct gccgcgccat ccgccacac 4260
ccccgcgca tcgccaggg cctggagcgc atcctgctgt gaggatccag atctgctgtg 4320
ccttctagtt gccagccatc tgttgtttgc cctcccccg tgcttctctt gacctggaa 4380
ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 4440
aggtgtcatt ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa 4500
gacaatagca ggcattgctg ggatgcggtg ggctctatgg gtaccaggt gctgaagaat 4560
tgacccggtt cctcctgggc cagaaagaag caggcacatc cccttctctg tgacacacc 4620
tgtccacgcc cctggttctt agttccagcc ccactcatag gacactcata gctcaggagg 4680
gctccgcctt caatcccacc cgctaaagta cttggagcgg tctctcctc cctcatcagc 4740
ccaccaaac aaacctagcc tccaagagtg ggaagaaatt aaagcaagat aggctattaa 4800
gtgcagaggg agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat 4860
tttaaggcca tcatggcctt aatcttccgc ttcctcgctc actgactcgc tgcgctcgg 4920
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 4980
atcaggggat aacgcaggaa agaactgtg agcaaaaggc cagcaaaagg ccaggaaccg 5040
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca 5100
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 5160
tccccctgga agtccctcg tgcgtctctc tgttccgacc ctgccgctta ccggatacct 5220
gtccgccttt ctccttctgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 5280
cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 5340
cgaccgctgc gccttatccg gtaactatcg tcttgagtc aaccggtaa gacacgactt 5400
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 5460
tacagagttc ttgaagtggg ggcctaacta cggtacact agaagaacag tatttggtat 5520
ctgcgctctg ctgaagccag ttacctctcg aaaaagagtt ggtagctctt gatccggcaa 5580
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 5640
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 5700
aaactcacgt taagggatct tggatcatgag attatcaaaa aggatcttca cctagatcct 5760
tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttggtctga 5820
cagttaccaaa tgcttaataa gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 5880
catagttgcc tgactcgggg gggggggggc ctgaggtctg cctcgtgaag aaggtgttgc 5940
tgactcatac caggcctgaa tcgcccac atccagccag aaagtgaggg agccacggtt 6000

```



```

gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttget ttgccacgga 6060
acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa aagttcgatt 6120
tattcaacaa agccgcccgc cgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa 6180
ttaaccaatt ctgattagaa aaactcatcg agcatcaa at gaaactgcaa tttattcata 6240
tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca 6300
ccgaggcagt tccataggat ggcaagatcc tggatcgggt ctgcgattcc gactcgtcca 6360
acatcaatac aacctattaa tttcccctcg tcaaaaataa gggtatcaag tgagaaatca 6420
ccatgagtga cgactgaatc cgggtgagaat ggcaaaagct tatgcatttc tttccagact 6480
tggtcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaaccgta 6540
ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa aggacaatta 6600
caaacaggaa tcgaatgcaa ccggcgagg aacactgcca gcgcatcaac aatattttca 6660
cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat cgcagtgggtg 6720
agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag aggcataaat 6780
tcgctcagcc agtttagtct gaccatctca tctgtaacat cattggcaac gctacctttg 6840
ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca 6900
cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc atccatgttg 6960
gaatttaatc gcggcctcga gcaagacgtt tcccggtgaa tatggctcat aacacccctt 7020
gtattactgt ttatgtaagc agacagtttt attgttcatt atgatataat tttatcttgt 7080
gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc ccattattga 7140
agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 7200
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 7260
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtc 7315

```

&lt;210&gt; 78

&lt;211&gt; 7498

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 78

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcatatgac ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tggcggttag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggcgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg ggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260

```

tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttaciaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	caccatgcgc	1920
gtgaaggaga	agtaccagca	cctgtggcgc	tggggctggc	gctggggcac	catgctgctg	1980
ggcatgctga	tgatctgcag	cgcaccgag	aagctgtggg	tgaccgtgta	ctacggcgctg	2040
cccgtgtgga	aggaggccac	caccaccctg	ctctgcgcca	gcgacgcca	ggcctacgac	2100
accgaggtgc	acaacgtgtg	ggccaccac	gcctgcgtgc	ccaccgaccc	caacccccag	2160
gaggtgggtgc	tggtcgacgt	gaccgagaac	ttcgacatgt	ggaagaacga	catggtggag	2220
cagatgcacg	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgtgaagctg	2280
acccccctgt	gcgtgagcct	gaagtgcacc	gacctgaaga	acgacaccaa	caccaacagc	2340
agcagcggcc	gcatgatcat	ggagaagggc	gagatcaagg	actgtagctt	cgacatcagc	2400
accagcatcc	gcggcaagg	gcagaaggag	tacgccttct	tctacaagct	ggacatcatc	2460
cccatcgaca	acgacaccac	cagctacagc	ctgaccagct	gcgaaactag	tgtgatcacc	2520
caggcctgcc	ccaaggtag	cttcgagccc	atccccaacc	actactgcgc	ccccgccggc	2580
ttcaccatcc	tgaagtgcaa	cgacaagacc	ttcgacggca	ccgggccctg	caccgacgtg	2640
agcaccgtgc	agtgcaccca	cggcatccgc	cccggtgtga	gcaccagct	gctgctgaac	2700
ggcagcctgg	ccgaggagga	ggtggtgatc	cgcagcgtcg	acttcaccga	caacgccaag	2760
accatcatcg	tgcaggtgga	caccacgggtg	gagatcgatt	gcaccgccc	caacaacaac	2820
acccgcaagc	gaattcgcat	ccagcgcggc	cccggccgcg	ccttcgtgac	catcggaag	2880
atcggcaaca	tgcgccagge	ccactgcgat	atcagccgcg	ccaagtggga	caacaccctg	2940
aagcagatcg	acagcaagct	gcgcgagcag	ttcggcaacg	acaagaccat	catcttcaag	3000
cagagcagcg	gcggcgaccc	cgagatcgtg	accacagct	tcaactgcgg	cggcgagttc	3060
ttctactgcg	acagcaccca	gctgttcaac	agcacctggt	tcaacagcac	ctggagcacc	3120
gagggcagca	acaacaccga	gggcagcgac	accatcaccc	tgccctgccg	catcaagcag	3180
atcatcaaca	tgtggcagaa	ggtgggcaag	gccatgtacg	cccccccat	cagcggccag	3240
atccgctgca	gcagcgatat	caccggcctg	ctgctgaccc	gcgacggcgg	caacagcaac	3300
aacgagagcg	agatcttccg	cctgggcggc	ggcgacatgc	gcgacaactg	gcgcagcgag	3360
ctgtacaagt	acaaggtagg	gaagatcgag	cccctgggcg	tggccccccac	caaggccaag	3420
cgcgcgctgg	tgcagcgca	gaagcgcgcc	gtgggcatcg	gcgcctgtgt	cctgggcttc	3480
ctgggcgccc	ccggcagcac	catgggcgcc	gccagcatga	ccctgaccgt	gcaggcccg	3540
cagctgctga	gcggcatcgt	gcagcagcag	aacaacctgc	tgcgcgccat	cgaggcccag	3600
cagcacctgc	tgcagctgac	cgtgtggggc	atcaagcagc	tgcaggcccg	catcctggcc	3660
gtggagcgct	acctgaagga	ccagcagctg	ctgggcatct	ggggctgcag	cggcaagctt	3720
atctgcacca	ccgcctgtgc	ctggaacgcc	agctggagca	acaagagcct	ggagcagatc	3780
tggaaaccaca	ccacctggat	ggagtgggac	cgcgagatca	acaactacac	cagcctgatc	3840
cacagcctga	tgcaggagag	ccagaaccag	caggagaaga	acgagcagga	gctgctggag	3900
ctggacaagt	gggccagcct	gtggaaactgg	ttcaacatca	ccaactggct	gtggtacatc	3960
aagctgttca	tcatgatcgt	gggcggcctg	gtgggcctgc	gcacgtgtgt	cgccgtgctg	4020
agcatcgtga	accgcgtgcg	ccagggtctac	agccccctga	gcttccagac	ccacctgccc	4080
atcccccgcg	gccccgaccg	ccccgagggc	atcgaggagg	agggcggcga	gcgcgaccgc	4140
gaccgcacga	tccgcctggg	gaacggcagc	ctggccctga	tctgggacga	cctgcgcagc	4200
ctgtgcctgt	tcagctacca	ccgcctgcgc	gacctgctgc	tgatcgtgac	ccgcacgtgc	4260
gagctgctgg	gccgcgcggg	ctgggaggcc	ctgaagtact	ggtggaacct	gctgcagtac	4320
tggagccagg	agctgaagaa	cagcgccgtg	agcctgctga	acgccaccgc	catcgccgtg	4380
gccgagggca	ccgaccgcgt	gatcgagggtg	gtgcagggcg	cctgccgcgc	catccgccac	4440
atcccccgcc	gcatccgcca	gggcctggag	cgcacccctg	tgtgaggatc	cagatctgct	4500
gtgccttcta	gttgccagcc	atctgttgtt	tgccccctcc	ccgtgccttc	cttgaccctg	4560

```

gaaggtgcc a ctcccactgt cctttcctaa taaaatgagg aaattgcac gcattgtctg 4620
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaagg ggaggattgg 4680
gaagacaata gcaggcatgc tggggatgcg gtgggctcta tgggtacca ggtgctgaag 4740
aattgacctg gttcctcctg ggccagaaag aagcaggcac atccccctt ctgtgacaca 4800
ccctgtccac gcccctggtt cttagtcca gcccactca taggacactc atagctcagg 4860
agggtccgc cttcaatccc acccgctaaa gtacttgag cggctctctc ctccctcac 4920
agcccaccaa accaaacctg gcctccaaga gtgggaagaa attaaagcaa gataggctat 4980
taagtgcaga gggagagaaa atgcctccaa catgtgagga agtaatgaga gaaatcatag 5040
aattttaagg ccatcatggc cttaatcttc cgcttcctcg ctactgact cgctgcgctc 5100
ggtcgttcgg ctgcggcgag cgggtatcag tcaactcaa gcggtataac ggttatccac 5160
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 5220
ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc cgccccctg acgagcatca 5280
caaaaatcga cgctcaagtc agaggtggcg aaaccgcaca ggactataaa gataccaggc 5340
gtttccccct ggaagctccc tcgtgcgctc tctgttccg accctgccgc ttaccggata 5400
cctgtccgcc tttctccctt cgggaagcgt ggcgtttct catagctcac gctgtaggta 5460
tctcagttcg gtgtaggctg ttcgctccaa gctgggctgt gtgcacgaac ccccggttca 5520
gcccagccgc tgcgccttat ccggtacta tcgtcttgag tccaaccgg taagacacga 5580
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 5640
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 5700
tatctgcgct ctgctgaagc cagttacctt cggaaaaga gttggtagct cttgatccgg 5760
caaacaacc accgctggta gcggtggttt ttttgttgc aagcagcaga ttacgcgcag 5820
aaaaaaagga tctcaagaag atcctttgat cttttctac gggctctgac ctcagtggaa 5880
cgaaaactca cgtaaggga ttttggctat gagattatca aaaaggatct tcacctagat 5940
ccttttaa at taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 6000
tgacagttac caatgcttaa tcagttaggc acctatctca gcgactctgc tatttcgttc 6060
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 6120
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 6180
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 6240
ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 6300
atztatcaa caaagccgcc gtcccgtaaa gtcagcgtaa tgctctgcca gtgttacaac 6360
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 6420
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatag aggagaaaac 6480
tcaccgaggc agttccatag gatggcaaga tcctggatc ggtctgcgat tccgactcgt 6540
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 6600
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 6660
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 6720
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6780
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 6840
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttccccgg gatcgagatg 6900
gtgagtaacc atgcatcac aggagtacgg ataaaatgct tgatggctcg aagaggcata 6960
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 7020
ttgccatggt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 7080
gcacctgatt gcccagacatt atcgcgagcc catttatacc catataaatc agcatccatg 7140
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 7200
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 7260
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccatattt 7320
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 7380
aataa caaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 7440
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 7498

```

&lt;210&gt; 79

&lt;211&gt; 7498

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 79

```
tgcgcgcttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtcccgctt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagc gtagtctgag cagtactcgt tgcgtccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga caccatgcgc 1920
gtgaaggaga agtaccagca cctgtggcgc tggggctggc gctggggcac catgctgctg 1980
ggcatgctga tgatctgcag cgccaccgag aagctgtggg tgaccgtgta ctacggcgtg 2040
cccgtgtgga aggaggccac caccaccctg ctctgcgcca gcgacgcaa ggcctacgac 2100
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 2160
gaggtggtgc tggctgcagc gaccgagaac ttcgacatgt ggaagaacga catggtggag 2220
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 2280
acccccctgt gcgtgagcct gaagtgcacc gacctgaaga acgacaccaa caccaacagc 2340
agcagcggcc gcatgatcat ggagaagggc gagatcaagg actgtagctt cgacatcagc 2400
accagcatcc gcggcaaggc gcagaaggag tacgccttct tctacaagct ggacatcatc 2460
cccatcgaca acgacaccac cagctacagc ctgaccagct gcgaaactag tgtgatcacc 2520
caggcctgcc ccaaggtagg cttcgagccc atccccacc actactgcgc ccccgccggc 2580
ttcaccatcc tgaagtgcaa cgacaagacc ttcgacggca ccggggcctg caccgacgtg 2640
agcaccgtgc agtgaccca cggcatccgc cccgtggtga gcaccagct gctgctgaac 2700
ggcagcctgg ccgaggagga ggtggtgatc cgcagcgtcg acttcaccga caacgccaag 2760
accatcatcg tgcaggtgga caccacggtg gagatcgatt gcacccgccc caacaacaac 2820
acccgcaagc gaattcgcat ccagcgcggc cccggccgcg ccttcgtgac catcggcaag 2880
atcggcaaca tgcgccaggc ccactgcaac atcagccgcg ccaagtggaa caacaccctg 2940
```



aagcagatcg acagcaagct gcgcgagcag ttcggcaaca acaagaccat catcttcaag 3000  
cagagcagcg gcggcgaccc cgagatcgtg acccacagct tcaactgcgg cggcgagttc 3060  
ttctactgca acagcaccca gctgttcaac agcacctggt tcaacagcac ctggagcacc 3120  
gagggcagca acaacaccga gggcagcgac accatcaccc tgccctgccg catcaagcag 3180  
atcatcaaca tgtggcagaa ggtgggcaag gccatgtacg cccccccat cagcggccag 3240  
atccgctgca gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caacagcaac 3300  
aacgagagcg agatcttccg cctgggcggc ggcgacatgc gcgacaactg gcgcagcgag 3360  
ctgtacaagt acaagggtgt gaagatcgag cccctgggcg tggccccccac caaggccaag 3420  
cgccgcgtgg tgcagcgcgga gaagcgcgcc gtgggcatcg gcgcccctgtt cctgggcttc 3480  
ctgggcggcg ccggcagcac catgggcggc gccagcatga ccctgaccgt gcaggcccgc 3540  
cagctgctga gcggcatcgt gcagcagcag aacaacctgc tgcgcgccat cgaggcccag 3600  
cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg catcctggcc 3660  
gtggagcgct acctgaagga ccagcagctg ctgggcatct ggggctgcag cggcaagctt 3720  
atctgcacca ccgccgtgcc ctggaacgcc agctggagca acaagagcct ggagcagatc 3780  
tggaaccaca ccacctggat ggagtgggac cgcgagatca acaactacac cagcctgatc 3840  
cacagcctga tgcaggagag ccagaaccag caggagaaga acgagcagga gctgctggag 3900  
ctggacaagt gggccagcct gtggaactgg ttcaacatca ccaactggct gtggtacatc 3960  
aagctgttca tcatgatcgt gggcggcctg gtgggcctgc gcacgtgtt cgccgtgctg 4020  
agcatcgtga accgcgtgcg ccagggtctac agccccctga gcttccagac ccacctgcc 4080  
atccccgcg gccccgaccg ccccgagggc atcgaggagg agggcggcga gcgcgaccgc 4140  
gaccgcacga tccgctggt gaacggcagc ctggccctga tctgggacga cctgcgcagc 4200  
ctgtgcctgt tcagctacca ccgctgcgc gacctgctgc tgatcgtgac ccgcatcgtg 4260  
gagctgctgg gccgcgcgg ctgggaggcc ctgaagtact ggtggaacct gctgcagtac 4320  
tgagaccagg agctgaagaa cagcgcctg agcctgctga acgccaccgc catcgccgtg 4380  
gccgaggga ccgaccgcgt gatcgagggt gtgcagggcg cctgccgcgc catccgccac 4440  
atccccgcc gcacccgcca gggcctggag cgcacccctgc tgtgaggatc cagatctgct 4500  
gtgccttcta gttgccagcc atctgttgtt tgccccctcc ccgtgccttc cttgacctg 4560  
gaaggtgcca ctccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg 4620  
agtaggtgtc attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg 4680  
gaagacaata gcaggcatgc tggggatgcg gtgggctcta tgggtacca ggtgctgaag 4740  
aattgaccgg gttcctcctg ggccagaaaag aagcaggcac atccccctt ctgtgacaca 4800  
ccctgtccac gcccctgggt cttagtcca gccccactca taggacactc atagctcagg 4860  
agggtccgc cttcaatccc acccgctaaa gtacttggag cggctctctc cttccctcatc 4920  
agcccaccaa accaaacct gcctccaaga gtgggaagaa attaaagcaa gataggctat 4980  
taagtgcaga gggagagaaa atgcctccaa catgtgagga agtaatgaga gaaatcatag 5040  
aattttaagg ccatcatggc cttaatcttc cgcttccctg ctcactgact cgctgcgctc 5100  
ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaag gcggtaatatc ggttatccac 5160  
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 5220  
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 5280  
caaaaatcga cgctcaagtc agagggtggc aaaccgcaca ggactataaa gataccaggc 5340  
gtttccccct ggaagctccc tcgtgcgctc tccgtgttcc accctgccgc ttaccggata 5400  
cctgtccgc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 5460  
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 5520  
gcccgaccgc tgcgccttat ccggtacta tegtcttgag tccaaccgg taagacacga 5580  
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgagg atgtaggcgg 5640  
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 5700  
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 5760  
caaaacaaacc accgctggta gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag 5820  
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 5880  
cgaaaactca cgttaaggga ttttgggtcat gagattatca aaaaggatct tcacctagat 5940  
ccttttaaat taaaaatgaa gtttttaaatc aatctaaagt atatatgagt aaacttggtc 6000  
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 6060  
atccatagtt gcctgactcg gggggggggg gcgctgagg ctgcctcgtg aagaagggtg 6120  
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 6180  
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 6240

```

ggaacggtct gcgttgctgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 6300
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 6360
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatte 6420
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 6480
tcaccgaggc agttccatag gatggcaaga tcctgggtatc ggtctgcatc tccgactcgt 6540
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgaagaa 6600
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 6660
acttggtcaa caggccagcc attacgctcg tcatcaaaat cactcgcac aaccaaaccc 6720
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctggt aaaaggacaa 6780
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatattt 6840
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg gatcgagtg 6900
gtgagtaacc atgcatcac aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6960
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 7020
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 7080
gcacctgatt gcccgcacat atcgcgagcc catttatacc catataaatc agcatccatg 7140
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 7200
cttgatttac tgtttatgta agcagacagt ttattgttc atgatgatat atttttatct 7260
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 7320
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 7380
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 7440
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 7498

```

&lt;210&gt; 80

&lt;211&gt; 5549

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 80

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320

```



```

ggatgggggtc ccatttatta tttaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga ccaccatggg 1920
cggcaagtgg agcaagagca gcgtgatcgg ctggcccgcg gtgcgcgagc gcatgcgcgc 1980
cgccgagccc gccgccgacg gcgtgggcgc cgtgagccgc gacctggaga agcacggcgc 2040
catcaccagc agcaacaccg ccgccaacaa cgccgcctgc gcctggctgg agggccagga 2100
ggaggaggag gtgggcttcc ccgtgacccc ccagggtccc ctgcgccccca tgacctacaa 2160
ggccgccgtg gacctgagcc acttccctgaa ggagaagggc ggcctggagg gcctgatcca 2220
cagccagcgc cgccaggaca tcctggacct gtggatctac cacaccagc gctacttccc 2280
cgactggcag aactacaccc ccggcccccg cgtgcgctac cccctgacct tcggctggtg 2340
ctacaagctg gtgcccgtgg agcccagaaa ggtggaggag gccaacaagg gcgagaacac 2400
cagcctgctg caccctgtga gcctgcacgg catggacgac ccgagcgcgc aggtgctgga 2460
gtggcgcttc gacagccgcc tggccttcca ccacgtggcc cgcgagctgc accccgagta 2520
cttcaagaac tgctgaggat ccagatctgc tgtgccttct agttgccagc catctgttgt 2580
ttgcccctcc ccggtgcctt ccttgacctt ggaaggtgcc actcccactg tcctttccta 2640
ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtg 2700
ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 2760
ggtgggctct atgggtaccc aggtgctgaa gaattgaccc ggttcctcct gggccagaaa 2820
gaagcaggca catccccttc tctgtgacac accctgtcca cgccctggtt tcttagttcc 2880
agccccactc ataggacact catagctcag gagggtccg ccttcaatcc caccgctaa 2940
agtacttggg gcggtctctc cctccctcat cagcccacca aaccaaacct agcctccaag 3000
agtgggaaga aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca 3060
acatgtgagg aagtaatgag agaaatcata gaattttaag gccatcatgg ccttaatctt 3120
ccgcttcttc gctcactgac tcgctgcgct cggctcgctc gctgcggcga gcggtatcag 3180
ctcactcaaa ggcggttaata cgtttatcca cagaatcagg ggataacgca ggaaagaaca 3240
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 3300
tccataggct ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 3360
gaaacccgac aggactataa agataccagg cgtttccccc tggaaagctc ctgctgcgct 3420
ctcctgttcc gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg 3480
tgccgcttcc tcatagctca cgtgttaggt atctcagttc ggtgtaggtc gttcgctcca 3540
agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggttaact 3600
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 3660
acaggattag cagagcagag tatgtaggcg gtgctacaga gttcttgaag tgggtggccta 3720
actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 3780
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcgggtggt 3840
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 3900
tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttgggtca 3960
tgagattatc aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agttttaaata 4020
caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg 4080
cacctatctc agcgatctgt ctatttctgt catccatagt tgctgactc gggggggggg 4140
ggcgtgagg tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc 4200
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 4260
agttggtgat tttgaacttt tgctttgcca cggaacggtc tgcgttgtcg ggaagatgcg 4320
tgatctgac cttcaactca gcaaaagttc gatttatcca acaaagccgc cgtcccgtca 4380
agtcagcgta atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc 4440
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 4500
aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 4560
atcctggtat cgggtctgca ttccgactcg tccaacatca atacaaccta ttaatttccc 4620

```

```

ctcgtcaaaa ataagggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 4680
gaatggcaaaa agcttatgca tttctttcca gacttggttca acaggccagc cattacgctc 4740
gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg cctgagcgag 4800
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg 4860
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 4920
ctggaatgct gttttcccgg ggatcgcatg ggtgagtaac catgcatcat caggagtacg 4980
gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat 5040
ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 5100
atcgggcttc ccatacaatc gatagattgt cgcacctgat tgcccagacat tatcgcgagc 5160
ccatttatac ccataataat cagcatccat gttggaattt aatcgcgggc tcgagcaaga 5220
cgtttcccgct tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag 5280
ttttattgtt catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga 5340
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400
gagcggatac atatttgaat gtatttagaa aaataaacia atagggttc cgcgcacatt 5460
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 5520
aaataggcgt atcacgagge ctttctgctc

```

&lt;210&gt; 81

&lt;211&gt; 5549

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 81

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc ttgcatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaaaaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatat gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgcct ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagcgt tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620

```

```

gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcaagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga ccaccatggg 1920
cggcaagtgg agcaagagca gcgtgatcgg ctggccccgc gtgcgcgagc gcatgcgcgc 1980
cgccgagccc gccgccgacg gcgtgggcgc cgtgagccgc gacctggaga agcacggcgc 2040
catcaccagc agcaacaccg ccgccaaaca cgccgcctgc gcctggctgg aggccagga 2100
ggaggaggag gtgggcttcg cagtaccgc acaggtggca ctgcgcgcaa tgacctaca 2160
ggccgccgtg gacctgagcc acttcttgaa ggagaagggc ggctggagg gcctgatcca 2220
cagccagcgc cgccaggaca tcctggacct gtggatctac cacaccagg gctacttccc 2280
cgactggcag aactacaccc ccggccccgc cgtgcgtac cccctgacct tcggctgggtg 2340
ctacaagctg gtgcccgtgg agcccagaca ggtggaggag gccaaacaagg gcgagaacac 2400
cagcctgctg caccctgtga gcctgcacgg catggcagca ccgagcgcg aggtgctgga 2460
gtggcgcttc gacagccgcc tggccttcca ccacgtggcc cgcgagctgc accccgagta 2520
cttcaagaac tgctgaggat ccagatctgc tgtgccttct agttgccagc catctgttgt 2580
ttgccccctc ccgctgcctt ccttgacct ggaagggtgc actcccactg tcctttccta 2640
ataaaatgag gaaattgcat cgcattgtct gtaggtgtt cattctattc tgggggggtg 2700
ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 2760
ggtgggctct atgggtaccc aggtgctgaa gaattgaccc ggttctcctt gggccagaaa 2820
gaagcaggca catccccctt tctgtgacac accctgtcca cggccctggt tcttagttcc 2880
agccccactc ataggacact catagctcag gagggctccg ccttcaatcc caccgccta 2940
agtacttgga gcggtctctc cctccctcat cagccacca aaccaaacct agcctccaag 3000
agtgggaaga aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca 3060
acatgtgagg aagtaatgag agaaatcata gaattttaag gccatcatgg ccttaatctt 3120
ccgcttcttc gctcactgac tcgctgcgct cggctcgttc gctgcggcga gcggtatcag 3180
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 3240
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 3300
tccataggct ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagagggtggc 3360
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgctgcgt 3420
ctcctgttcc gacctgcgcg cttaccggat acctgtccgc ctttctccct tcgggaagcg 3480
tggcgcttcc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttogetcca 3540
agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggtaact 3600
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 3660
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta 3720
actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 3780
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt 3840
tttttggttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 3900
tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca 3960
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 4020
caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 4080
cacctatctc agcgatctgt ctatttcgtt catccatagt tgctgactc gggggggggg 4140
ggcgctgagg tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc 4200
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 4260
agttgggtgat tttgaacttt tgctttgcca cggaacgggtc tgcgttgctg ggaagatgcg 4320
tgatctgate cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca 4380
agtcagcgta atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc 4440
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 4500
aaaaagccgt ttctgtaatg aaggagaaaa ctaccgagg cagttccata ggatggcaag 4560
atcctgggat cggtctgcga ttccgactcg tccaacatca atacaacctt ttaatttccc 4620
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 4680
gaatggcaaa agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc 4740
gtcatcaaaa tcaactcgcat caaccaaacc gttattcatt cgtgatttgc cctgagcgag 4800
acgaaatacg cgatcgctgt taaaaggaca attacaaca ggaatcgaat gcaaccggcg 4860
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 4920

```

```

ctggaatgct gttttcccg ggcacgcagt ggtgagtaac catgcatcat caggagtacg 4980
gataaaatgc ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgaccat 5040
ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 5100
atcgggcttc ccatacaatc gatagattgt cgcacctgat tgcccgcacat tatcgcgagc 5160
ccatttatac ccataataat cagcatccat gttggaattt aatcgcggcc tcgagcaaga 5220
cgtttcccg tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 5280
ttttattgtt catgatgata ttttttatac ttgtgcaatg taacatcaga gattttgaga 5340
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400
gagcggatgc atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt 5460
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 5520
aaataggcgt atcacgaggc cttttcgtc 5549

```

&lt;210&gt; 82

&lt;211&gt; 5549

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 82

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga ccaccatggg 1920

```



cggcaagtgg	agcaagagca	gcgtgatcgg	ctggcccggc	gtgcgcgagc	gcatgcgcgc	1980
cgccgagccc	gccgccgacg	gcgtggggcg	cgtgagccgc	gacctggaga	agcacggcgc	2040
catcaccagc	agcaacaccg	ccgccaaaca	cgccgcctgc	gcctggctgg	aggcccagga	2100
ggaggaggag	gtgggcttcc	ccgtgacccc	ccaggtgccc	ctgcgcccc	tgacctacaa	2160
ggccgcctg	gacctgagcc	acttcctgaa	ggagaagggc	ggcctggagg	gcctgatcca	2220
cagccagcgc	cgccaggaca	tcctggacct	gtggatctac	cacacccagg	gctacttccc	2280
cgactggcag	aactacaccc	ccggccccgg	cgtgcgctac	cccctgacct	tcggctggtg	2340
ctacaagctg	gtgcccgtgg	agccccgaca	ggtggaggag	gccaacaagg	gcgagaacac	2400
cagcctgctg	caccccgtga	gcctgcacgg	catggcagca	cccagcgcgc	aggtgctgga	2460
gtggcgcttc	gacagccgcg	tggccttcca	ccacgtggcc	cgcgagctgc	accccagagta	2520
cttcaagaac	tgctgaggat	ccagatctgc	tgtgccttct	agttgccagc	catctgttgt	2580
ttgcccctcc	cccgtgcctt	ccttgacctt	ggaaggtgcc	actcccactg	tcctttccta	2640
ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctatct	tggggggtgg	2700
ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	2760
ggtgggctct	atgggtaccc	aggtgctgaa	gaattgaccc	ggttcctcct	gggccagaaa	2820
gaagcaggca	catccccttc	tctgtgacac	accctgtcca	cgcccctggc	tcttagttcc	2880
agccccactc	ataggacact	catagctcag	gagggctccg	ccttcaatcc	caccgcctaa	2940
agtacttgga	gcggtctctc	cctccctcat	cagcccacca	aaccaaacct	agcctccaag	3000
agtgggaaga	aattaaagca	agataggcta	ttaagtgcag	aggagagaaa	aatgcctcca	3060
acatgtgagg	aagtaatgag	agaaatcata	gaattttaag	gccatcatgg	ccttaatctt	3120
ccgcttcctc	gctcactgac	tcgctgcgct	cggctcgttc	gctgcggcga	gcggtatcag	3180
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	3240
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	3300
tccataggct	ccgccccctt	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	3360
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaagctcc	ctcgtgcgct	3420
ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	3480
tggcgcttcc	tcatagtcca	cgtgttaggt	atctcagttc	ggtgtaggtc	gttcgctcca	3540
agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	ctgcgcctta	tccggtaact	3600
atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	3660
acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtagccta	3720
actacggcta	cactagaaga	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	3780
tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	caccgctggc	agcgggtggt	3840
tttttgtttg	caagcagcag	attacgcgca	gaaaaaaaag	atctcaagaa	gatcctttga	3900
tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	acgttaaggg	attttggcca	3960
tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaatga	agtttttaaa	4020
caatctaaag	tatatatgag	taaacttggc	ctgacagtta	ccaatgctta	atcagtgagg	4080
cacctatctc	agcgatctgt	ctatctcggt	catccatagt	tgcttgactc	gggggggggg	4140
ggcgctgagg	tctgcctcgt	gaagaagggt	ttgctgactc	ataccaggcc	tgaatcgccc	4200
catcatccag	ccagaaagtg	aggagagccac	ggttgatgag	agctttgttg	taggtggacc	4260
agttgggtgat	tttgaacttt	tgtcttgcca	cggaacgggc	tgctgtgtcg	ggaagatgcg	4320
tgatctgate	cttcaactca	gcaaaaagtt	gatttattca	acaaagccgc	cgccccgtca	4380
agtcagcgta	atgctctgcc	agtgttacia	ccaattaaac	aattctgatt	agaaaaactc	4440
atcgagcacc	aaatgaaact	gcaatttatt	catatcagga	ttatcaatac	catatttttg	4500
aaaaagccgt	ttctgtaatg	aaggagaaaa	ctcaccgagg	cagttccata	ggatggcaag	4560
atcctgggat	cggctctgcg	ttccgactcg	tccaacatca	atacaacct	ttaatttccc	4620
ctcgtcaaaa	ataagggtat	caagtgagaa	atcaccatga	gtgacgactg	aatccgggtg	4680
gaatggcaaa	agcttatgca	tttctttcca	gacttggttc	acaggccagc	cattacgctc	4740
gtcatcaaaa	tactcgcac	caaccaaacc	gttattcatt	cgtgattgcg	cctgagcgag	4800
acgaaatacg	cgatcgctgt	taaaaggaca	attacaaaca	ggaatcgaat	gcaaccggcg	4860
caggaaacact	gccagcgcac	caacaatatt	ttcacctgaa	tcaggatatt	cttctaatac	4920
ctggaatgct	gttttcccgg	ggatcgaggt	ggtgagtaac	catgcatcat	caggagtacg	4980
gataaaatgc	ttgatgggtc	gaagaggcat	aaattccgtc	agccagttta	gtctgaccat	5040
ctcatctgta	acatcattgg	caacgctacc	tttgccatgt	ttcagaaaca	actctggcgc	5100
atcgggcttc	ccatacaatc	gatagattgt	cgcacctgat	tgcccagacat	tatcgcgagc	5160
ccattttatac	ccatataaat	cagcatccat	gttggaattt	aatcgcgggc	tcgagcaaga	5220

```

cgtttcccgt tgaatatggc tcataacacc ccttgtatta ctgtttatgt aagcagacag 5280
ttttattggt catgatgata tattttttatc ttgtgcaatg taacatcaga gatttttgaga 5340
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 5400
gagcggatac atatttgaat gtatttagaa aaataaacia ataggggttc cgcgacatt 5460
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 5520
aaataggcgt atcacgaggc cctttcgtc 5549

```

&lt;210&gt; 83

&lt;211&gt; 5549

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 83

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc battaccatg gtgatgcggg ttgggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaaaaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatth gaacgcggat 1020
tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggth tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacgc gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860
tgcaagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga ccaccatggg 1920
cggcaagtgg agcaagagca gcgtgatcgg ctggccccgc gtgcgcgagc gcatgcgcgc 1980
cgccgagccc gccgccgacg gcgtggggcg cgtgagccgc gacctggaga agcacggcg 2040
catcaccagc agcaaacacc ccgccaaaaa cgccgcctgc gcctggctgg aggccagga 2100
ggaggaggag gtgggcttcg cagtgaaccg acaggtggca ctgcgcgcaa tgacctaaa 2160
ggccgcgctg gacctgagcc acttcctgaa ggagaagggc ggcctggagg gcctgatcca 2220

```



cagccagcgc	cgccaggaca	tcctggacct	gtggatctac	cacacccagg	gctacttccc	2280
cgactggcag	aactacaccc	cgggccccgg	cgtgcgctac	cccctgacct	tcggctggtg	2340
ctacaagctg	gtgcccgtgg	agcccgacaa	ggtaggaggag	gccaacaagg	gcgagaacac	2400
cagcctgctg	caccccgtga	gcctgcacgg	catggacgac	cccagagcgc	aggtgctgga	2460
gtggcgcttc	gacagccgcc	tggccttcca	ccacgtggcc	cgcgagctgc	accccagta	2520
cttcaagaac	tgctgaggat	ccagatctgc	tgtgccttct	agttgccagc	catctgttgt	2580
ttgccccctc	cccgtgcctt	ccttgacct	ggaagggtgc	actcccactg	tcctttccta	2640
ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	tggggggtgg	2700
ggtagggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	2760
ggtagggctct	atgggtaccc	aggtgctgaa	gaattgaccc	ggttcctcct	gggccagaaa	2820
gaagcaggca	catccccttc	tctgtgacac	accctgtcca	cgccccctgg	tcttagttcc	2880
agccccactc	ataggacact	catagctcag	gagggctccg	ccttcaatcc	caccgcctaa	2940
agtacttgga	gcggtctctc	cctccctcat	cagcccacca	aaccaaacct	agcctccaag	3000
agtggggaaga	aattaaagca	agataggcta	ttaagtgcag	aggagagaaa	aatgcctcca	3060
acatgtgagg	aagtaatgag	agaaatcata	gaattttaag	gccatcatgg	ccttaatctt	3120
ccgcttcctc	gctcactgac	tcgctgcgct	cggctcgttcg	gctgcggcga	gcggtatcag	3180
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	3240
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgctg	ctggcgtttt	3300
tccataggct	ccgccccctc	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	3360
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaaagctcc	ctcgtgcgct	3420
ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	3480
tggcgctttc	tcatagctca	cgtgttaggt	atctcagttc	ggtaggtg	gttcgctcca	3540
agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	ctgcgcctta	tccggtaact	3600
atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	3660
acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtagccta	3720
actacggcta	cactagaaga	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	3780
tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	caccgctggg	agcggtaggt	3840
tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gatcctttga	3900
tcttttctac	ggggtctgac	gtcagtgga	acgaaaactc	acgttaaggg	attttggtea	3960
tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaatga	agttttaaat	4020
caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	atcagtgagg	4080
cacctatctc	agcgtctgt	ctatttcggt	catccatagt	tgctgactc	gggggggggg	4140
ggcgctgagg	tctgcctcgt	gaagaagggt	ttgctgactc	ataccaggcc	tgaatcgccc	4200
catcatccag	ccagaaagtg	aggagaccac	ggtagatgag	agctttgttg	taggtggacc	4260
agttgggtgat	tttgaacttt	tgttttgcca	cggaaacggtc	tgctgtgtcg	ggaagatgcg	4320
tgatctgac	cttcaactca	gcaaaagttc	gatttattca	acaaagccgc	cgtcccgtca	4380
agtcagcgta	atgctctgcc	agtgttataa	ccaattaaac	aattctgatt	agaaaaactc	4440
atcgagcatc	aaatgaaact	gcaatttatt	catatcagga	ttatcaatac	catatttttg	4500
aaaaagccgt	ttctgtaatg	aaggagaaaa	ctcaccgagg	cagttccata	ggatggcaag	4560
atcctgggat	cggctctgcga	ttccgactcg	tccaacatca	atacaacctt	ttaatttccc	4620
ctcgtcaaaa	ataaggttat	caagtgaaga	atcaccatga	gtgacgactg	aatccggtga	4680
gaatggcaaa	agcttatgca	tttctttcca	gacttggttc	acaggccagc	cattacgctc	4740
gtcatcaaaa	tcactcgcat	caaccaaacc	gttattcatt	cgtgattgcy	cctgagcgag	4800
acgaaatacg	cgatcgctgt	taaaaggaca	attacaaaca	ggaatcgaat	gcaaccggcg	4860
caggaaact	gccagcgcat	caacaatatt	ttcacctgaa	tcaggatatt	cttctaatac	4920
ctggaatgct	gttttcccgg	ggatcgaggt	ggtagtaaac	catgcatcat	caggagtacg	4980
gataaaatgc	ttgatggctg	gaagaggcat	aaattccgct	agccagttaa	gtctgaccat	5040
ctcatctgta	acatcattgg	caacgctacc	tttgccatgt	ttcagaaaca	actctggcgc	5100
atcgggcttc	ccatacaatc	gatagattgt	cgcacctgat	tgcccagacat	tatcgcgagc	5160
ccattttatac	ccatataaat	cagcatccat	gttggaattt	aatcgcgggc	tcgagcaaga	5220
cgtttcccgt	tgaatatggc	tcataacacc	ccttgattta	ctgtttatgt	aagcagacag	5280
ttttattgtt	catgatgata	tattttttatc	ttgtgcaatg	taacatcaga	gattttgaga	5340
cacaacgtgg	ctttcccccc	cccccattha	ttgaagcatt	tatcaggggt	attgtctcat	5400
gagcggatag	atatttgaat	gtatbtagaa	aaataaacia	ataggggttc	cgcgcacatt	5460
tccccgaaaa	gtgccacctg	acgtctaaga	aaccattatt	atcatgacat	taacctataa	5520

aaataggcgt atcacgaggc cctttcgtc

5549

&lt;210&gt; 84

&lt;211&gt; 6609

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 84

```
tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat ggggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tggggcacca 1980
tgctgctggg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accgtgtact 2040
acggcgtgcc cgtgtggcgc gaggccacca ccacctgtt ctgcgccagc gacgccaagg 2100
cctacgacac cgaggtgcac aacgtgtggg ccaccacgc ctgcgtgccc accgacccca 2160
acccccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220
tggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280
tgaagctgac cccctgtgc gtgaccctga actgcaccaa cctgaacatc accaagaaca 2340
ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400
gcttctacat caccaccagc atccgcaaca aggtgaagaa ggagtacgcc ctgttcaacc 2460
gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccgctg atcagctgca 2520
```

```

acaccagcgt gatcaccag gcctgcccc aggtgagctt ccagcccatc cccatccact 2580
actgcgtgcc cgccggcttc gccatcctga agtgcaacaa caagaccttc aacggcagcg 2640
gcccctgcac caacgtgagc accgtgcagt gcacccacgg catccgcccc gtggtgagca 2700
cccagctgct gctgaacggc agcctggccg aggaggacat cgtgatccgc agcgaggact 2760
tcaccgacaa cgtgaagacc atcatcgtgc agctgaacga gagcgtggtg atcaactgca 2820
cccgcccaaa caacaacacc cgcgagcgcc tgagcatcgg ccccgccgcg gccttctacg 2880
cccgccgcaa catcatcggc gacatccgcc aggccactg caacatcagc cgcgccaagt 2940
ggaacaacac cctgcagcag atcgtgatca agctgcgcga gaagttccgc aacaagacca 3000
tcgcttcaa ccagagcagc ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg 3060
gcggcgagtt cttctactgc aacaccgccc agctgttcaa cagcacctgg aacgtggccg 3120
gcggcaccaa cggcaccgag ggcaacgaca tcatcaccct gcagtgccgc atcaagcaga 3180
tcatcaacat gtggcagaag gtgggcaagg ccatgtacgc cccccccatc accggccaga 3240
tccgctgcag cagcaacatc accggcctgc tgctgacctg cgacggcggc aacagcaccg 3300
agaccgagac cgagatcttc cgccccggcg gcggcgacat gcgcgacaaac tggcgagcgc 3360
agctgtacaa gtacaagggt gtgcgcacgc agcccatcgg cgtggccccc acccgcgcca 3420
agcttaccgt gcaggcccg cctgctgctga gcggcatcgt gcagcagcag aacaacctgc 3480
tgcgcgccat cgaggcccag cagaacatgc tgcgcctgac cgtgtggggc atcaagcagc 3540
tgaggcccg cgtgctggcc ctggagcgct acctgcgcga ccagcagctc gagtgaggat 3600
ccagatctgc tgtgccttct agttgccagc catctgttgt ttgccccctc cccgtgcctt 3660
ccttgacctt ggaagggtgc actcccactg tcccttccca ataaaatgag gaaattgcat 3720
cgcatctgtc gagtaggtgt cattctatct tggggggtgg ggtggggcag gacagcaagg 3780
gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atgggtacct 3840
aggtgctgaa gaattgacct ggttccctct gggccagaaa gaagcaggca catccccctc 3900
tctgtgacac acctgtcca cgccccgtgt tcttagttcc agccccactc ataggacact 3960
catagctcag gagggtcccg ccttcaatcc caccgcctaa agtacttgga gcggtctctc 4020
cctccctcat cagcccacca aaccaaacct agcctccaag agtgggaaga aattaaagca 4080
agataggcta ttaagtgcag agggagagaa aatgcctcca acatgtgagg aagtaatgag 4140
agaaatcata gaattttaag gccatcatgg ccttaatctt ccgcttccct gctcactgac 4200
tcgctgcgct cggctcgttcg gctgcggcga gcggtatcag ctcaactcaa ggccgtaata 4260
cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 4320
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccctc 4380
gacgagcatc acaaaaatcg acgctcaagt cagagggtggc gaaacccgac aggactataa 4440
agataccagg cgtttccccc tggaagctcc ctgctgcgct ctctgttcc gacctgccg 4500
cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttct tcatagctca 4560
cgctgtaggt atctcagttc ggtgtaggtc gtctcgctcca agctgggctg tgtgcacgaa 4620
cccccgctc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacctg 4680
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 4740
tatgtaggcg gtgctacaga gttcttgaag tggtaggcta actacggcta cactagaaga 4800
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 4860
tcttgatccg gcaaacaaac caccgctggc agcgggtggt tttttgtttg caagcagcag 4920
attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 4980
gctcagtggg acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc 5040
ttcacctaga tctttttaa ttaaaaatga agttttaaat caatctaaag tatatatgag 5100
taaacttggc ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt 5160
ctatttcgtt catccatagt tgccctgactc gggggggggg ggcgctgagg tctgcctcgt 5220
gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag ccagaaagtg 5280
aggagaccac gggtgatgag agctttgttg taggtggacc agttggtgat tttgaacttt 5340
tgctttgcca cggaacggtc tgcgttgctg ggaagatgcg tgatctgate cttcaactca 5400
gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta atgctctgcc 5460
agtgttacia ccaattaacc aattctgatt agaaaaactc atcgagcatc aaatgaaact 5520
gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt ttctgtaatg 5580
aaggagaaaa ctcccgagg cagttccata ggatggcaag atcctgggat cgggtctgca 5640
ttccgactcg tccaacatca atacaacctt ttaatttccc ctgctcaaaa ataaggttat 5700
caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa agcttatgca 5760
tttctttcca gacttgttca acaggccagc cattacgctc gtcataaaaa tcaactcgcat 5820

```



```

caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt 5880
taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgcat 5940
caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct gttttcccg 6000
ggatcgcgagt ggtgagtaac catgcatcat caggagtacg gataaaatgc ttgatgggtcg 6060
gaagaggcat aaattccgct agccagttta gtctgacctat ctcattctgta acatcattgg 6120
caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc 6180
gatagattgt cgcacctgat tgcccagacat tatcgcgagc ccattttatac ccatataaat 6240
cagcatccat gttggaattt aatcgcgggc tcgagcaaga cgtttcccg 6300
tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattgtt catgatgata 6360
tattttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg cttttcccc 6420
ccccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 6480
gtattttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 6540
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 6600
cctttcgctc 6609

```

&lt;210&gt; 85

&lt;211&gt; 6783

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 85

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacac aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaaaaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgcc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcaag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920  
 ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tggggcacca 1980  
 tgctgctggg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accgtgtact 2040  
 acggcgtgcc cgtgtggcgc gaggccacca ccacctgtt ctgcgccagc gacgccaagg 2100  
 cctacgacac cgaggtgcac aacgtgtggg ccaccacgc ctgcgtgccc accgaccca 2160  
 acccccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220  
 tgggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280  
 tgaagctgac cccctgtgc gtgacctga actgcaccaa cctgaacatc accaagaaca 2340  
 ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400  
 gcttctacat caccaccagc atccgcaaca aggtgaagaa ggagtacgcc ctgttcaacc 2460  
 gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccgcctg atcagctgca 2520  
 acaccagcgt gatcaccag gcctgcccc aggtgagctt ccagcccatc cccatccact 2580  
 actgcgtgcc cgcgggttc gccatcctga agtgcaacaa caagacctc aacggcagcg 2640  
 gcccctgcac caacgtgagc accgtgcagt gcaccacgg catccgccc gtggtgagca 2700  
 ccagctgct gctgaacggc agcctggccg aggaggacat cgtgatccgc agcgaggact 2760  
 tcaccgacaa cgtgaagacc atcatcgtgc agctgaacga gagcgtggtg atcaactgca 2820  
 cccgccccaa caacaacacc cgcgagcgcc tgagcatcgg ccccgccgc gccttctacg 2880  
 cccgcccga catcatcggc gacatccgcc aggccactg caacatcagc cgcgccaagt 2940  
 ggaacaacac cctgcagcag atcgtgatca agctgcgcga gaagttccgc aacaagacca 3000  
 tcgcttcaa ccagagcagc ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg 3060  
 gcggcgagtt cttctactgc aacaccgccc agctgttcaa cagcacctgg aacgtggccg 3120  
 gcggcaccaa cggcaccgag ggcaacgaca tcataccct gcagtccgc atcaagcaga 3180  
 tcatacaat gtggcagaag gtgggcaagg ccatgtacgc ccccccatc accggccaga 3240  
 tccgctgcag cagcaacatc accggcctgc tgctgacctg cgacggcggc aacagcaccg 3300  
 agaccgagac cgagatcttc cgcgccggcg gcggcgacat gcgcgacaac tggcgagcg 3360  
 agctgtacaa gtacaagggt gtgcgcacgc agcccatcgg cgtggcccc acccgcgcca 3420  
 agcttaccgt gcaggcccg ctgctgctga gcggcatcgt gcagcagcag aacaacctgc 3480  
 tgccgcccac cgaggcccag cagaacatgc tgccctgac cgtgtggggc atcaagcagc 3540  
 tgcaggcccg cgtgctggcc ctggagcgt acctgcgcga ccagcagctc gaggacatct 3600  
 ggaacaacat gacctggatg gagtgggagc gcgagatcga caactacacc gactacatct 3660  
 acgacctgct ggagaagagc cagaccagc aggagaagaa cgagaaggag ctgctggagc 3720  
 tggacaagtg ggccagcctg tggaaactggt tcgacatcac caactgggtga ggatccagat 3780  
 ctgctgtgcc ttctagtgtc cagccatctg ttgtttgccc cteccccgtg ccttccttga 3840  
 ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcat 3900  
 gtctgagtag gtgtcattct attctggggg gtgggggtgg gcaggacagc aagggggagg 3960  
 attgggaaga caatagcagg catgctgggg atgcgggtgg ctctatgggt acccaggtgc 4020  
 tgaagaattg acccggttcc tccctgggcca gaaagaagca ggcacatccc cttctctgtg 4080  
 acacacctg tccacgccc tggttcttag ttccagccc actcatagga cactcatagc 4140  
 tcaggagggc tccgccttca atcccacccg ctaaagtact tggagcggtc tctccctccc 4200  
 tcatacagccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4260  
 gctattaagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4320  
 catagaattt taaggccatc atggccttaa tcttccgctt cctcgctcac tgactcgctg 4380  
 cgctcggtcg ttccgctgcg gcgagcggta tcagctcact caaaggcggc aatacggtta 4440  
 tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 4500  
 aggaaccgta aaaaggccgc gttgctggcg tttttccata ggcctccgcc ccctgacgag 4560  
 catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 4620  
 caggcgtttc cccctggaag ctccctcgct cgctctctg ttccgacct gccgcttacc 4680  
 ggatacctgt ccgcctttct ccttccggga agcgtggcgc tttctcatag ctacagctgt 4740  
 aggtatctca gttcgggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 4800  
 gttcagccc accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 4860  
 cagcacttat cgcactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 4920  
 ggcggtgcta cagagtctt gaagtgggtg cctaactacg gctacactag aagaacagta 4980  
 tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 5040

```

tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg 5100
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag 5160
tggaacgaaa actcacgtta agggattttg gtcattgagat tatcaaaaag gatcttcacc 5220
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 5280
tggctctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 5340
cgttcatcca tagttgcctg actcgggggg gggggggcgt gaggtctgcc tcgtgaagaa 5400
ggtgttgctg actcatacca ggccctgaatc gcccctcatc ccagccagaa agtgaggag 5460
ccacgggtga tgagagcttt gttgtagggt gaccagtgtg tgattttgaa cttttgcttt 5520
gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa 5580
gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt 5640
acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt 5700
tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag 5760
aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcgggtc gcgattccga 5820
ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg 5880
agaaatcacc atgagtgcg actgaatccg gtgagaatgg caaaagctta tgcattttct 5940
tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc gcatcaacca 6000
aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaaag 6060
gacaattaca aacaggaatc gaatgcaacc ggcgaggaa cactgccagc gcatcaacaa 6120
tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg 6180
cagtggtag taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag 6240
gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc 6300
tacctttgcc atgtttcaga aacaactctg gcgcacggg cttcccatac aatcgataga 6360
ttgtcgcacc tgattgccc acattatcgc gagcccatth ataccatat aaatcagcat 6420
ccatgttgga atttaatcgc ggccctcgagc aagacgtttc ccgttgaata tggctcataa 6480
cacccttgt attactgtt atgtaagcag acagttttat tgttcattgat gatataattt 6540
tatcttgtgc aatgtaacat cagagatttt gagacacaac gtggctttcc ccccccccc 6600
attattgaag catttatcag ggttattgtc tcatgagcgg atacataatt gaatgtattt 6660
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct 6720
aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg aggcctttc 6780
gtc

```

&lt;210&gt; 86

&lt;211&gt; 6861

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 86

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgbct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatagcgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780

```



```

cgtcaatggg agtttgTTTT ggcacaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgTTta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagTTTTt attaaacata gcggtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggtgagc tccacatccg agccctgggtc ccatgcctcc 1500
agcggtctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgcgtcccg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat ggggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tggggcacca 1980
tgcgtcgtgg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accgtgtact 2040
acggcggtgcc cgtgtggcgc gagggcacca ccacctgtt ctgcgccagc gacgccaagg 2100
cctacgacac cgaggtgcac aacgtgtggg ccaccacgc ctgcgtgccc accgacccca 2160
acccccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220
tgggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280
tgaagctgac cccctgtgc gtgacctga actgcaccaa cctgaacatc accaagaaca 2340
ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400
gcttctacat caccaccagc atccgcaaca aggtgaagaa ggagtacgcc ctgttcaacc 2460
gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccgcctg atcagctgca 2520
acaccagcgt gatcaccag gcctgcccc aagtgagctt ccagcccatc cccatccact 2580
actgcgtgcc cgcggcttc gccatcctga agtgcaacaa caagacctc aacggcagcg 2640
gcccctgcac caacgtgagc accgtgcagt gcaccacgg catccgcccc gtggtgagca 2700
cccagctgct gctgaacggc agcctggcg aggaggacat cgtgatccgc agcgaggact 2760
tcaccgacaa cgtgaagacc atcatcgtgc agctgaacga gagcgtgggtg atcaactgca 2820
cccgcccaaa caacaacacc cgcgagcgcc tgagcatcgg ccccgggcgc gccttctacg 2880
cccgccgcaa catcatcggc gacatccgc agggccactg caacatcagc cgcgccaagt 2940
ggaacaacac cctgcagcag atcgtgatca agctgcgcga gaagtccgc aacaagacca 3000
tcgccttcaa ccagagcagc ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg 3060
gcggcgagtt cttctactgc aacaccgccc agctgttcaa cagcacctgg aacgtggccg 3120
gcggcaccaa cggcaccgag ggcaacgaca tcatcaccct gcagtgcgc atcaagcaga 3180
tcatcaacat gtggcagaag gtgggcaagg ccatgtacgc ccccccatc accggccaga 3240
tccgctgcag cagcaacatc accggcctgc tgcgtacccg cgacggcggc aacagcaccg 3300
agaccgagac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac tggcgagcg 3360
agctgtacaa gtacaagggt gtgcgcacat agcccatcgg cgtggcccc acccgcgcca 3420
agcttaccgt gcaggcccc ctgctgctga gcggcatcgt gcagcagcag aacaacctgc 3480
tgcgcgccat cgaggcccc agaacatgc tgcgcctgac cgtgtggggc atcaagcagc 3540
tgcaggcccc cgtgctggcc ctggagcgt acctgcgcga ccagcagctc gaggacatct 3600
ggaacaacat gacctggatg gagtgggagc gcgagatcga caactacacc gactacatct 3660
acgacctgct ggagaagagc cagaccagc aggagaagaa cgagaaggag ctgctggagc 3720
tggacaagtg ggccagcctg tggaaactgg tgcacatcac caactggctg tggtagatcc 3780
gcctgttcat catgatcgtg ggcggcctga tggcctgcg catcgtgttc gccgtgctga 3840
gcatctgagg atccagatct gctgtgcct ctagttgcca gccatctgtt gtttgccct 3900
ccccggtgcc ttccttgacc ctggaagggt ccactcccac tgccttttc taataaaaatg 3960
aggaaattgc atcgcatgt ctgagtaggt gtcattctat tctggggggg ggggtggggc 4020
aggacagcaa gggggaggat tgggaagaca atagcaggca tgcgtgggat gcggtgggct 4080

```

ctatgggtac ccaggtgctg aagaattgac ccggttcttc ctggggccaga aagaagcagg 4140  
cacatcccct tctctgtgac acaccctgtc cagcccccctg gttcttagtt ccagccccac 4200  
tcataggaca etcatagctc aggagggctc cgccttcaat cccacccgct aaagtacttg 4260  
gagcgggtctc tccctccctc atcagcccac caaaccaaac ctagectcca agagtgggaa 4320  
gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc caacatgtga 4380  
ggaagtaatg agagaaatca tagaatttta aggccatcat ggccttaate ttccgcttcc 4440  
tcgctcactg actcgtgctg ctcggtcggt cggtgctggc gagcgggtatc agctcactca 4500  
aaggcggtaa tacgggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 4560  
aaaggccagc aaaaggccag gaaccgtaaa aaggccgctg tgctggcggt tttccatagg 4620  
ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg gcgaaaccgc 4680  
acaggactat aaagatacca ggcgtttccc cctggaagct cctcgtgctg ctctcctggt 4740  
ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt 4800  
tctcatagct cagcgtgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 4860  
tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt 4920  
gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 4980  
agcagagcga ggtatgtagg cgggtgtaca gagttcttga agtgggtggc taactacggc 5040  
tacactagaa gaacagtatt tgggtatctgc gctctgctga agccagttac cttcggaaaa 5100  
agagtggta gctcttgatc cggcaaacaa accaccgctg gtagcgggtg tttttttgtt 5160  
tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct 5220  
acggggctctg acgctcagt gaacgaaaac tcacgttaag ggatttttgt catgagatta 5280  
tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 5340  
agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc 5400  
tcagcgatct gtctatctcg ttcattccata gttgcctgac tcgggggggg ggggcgctga 5460  
ggtctgcctc gtgaagaagg tgggtgtgac tcataccagg cctgaatcgc cccatcatcc 5520  
agccagaaaag tgaggaggcc acggttgatg agagctttgt tgtagggtga ccagttggtg 5580  
atcttgaaact tttgctttgc cacggaacgg tctgcgttgt cgggaagatg cgtgatctga 5640  
tccttcaact cagcaaaagt tcgatttatt caacaaagcc gccgtcccgt caagtcagcg 5700  
taatgctctg ccagtgttac aaccaattaa ccaattctga ttagaaaaac tcatcgagca 5760  
tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatctt tgaaaaagcc 5820  
gtttctgtaa tgaaggagaa aactcaccga ggcagttcca taggatggca agatcctggc 5880  
atcggtctgc gattccgact cgtccaacat caatacaacc tattaatttc cctcgtcaa 5940  
aaataagggt atcaagtgag aaatcaccat gagtgcgac tgaatccggt gagaatggca 6000  
aaagcttatg catctctttc cagacttggt caacaggcca gccattacgc tcgtcatcaa 6060  
aatcactcgc atcaacccaa ccgttattca ttcgtgattg cgcctgagcg agacgaaata 6120  
cgcgatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca 6180  
ctgccagcgc atcaacaata ttttcacctg aatcaggata ttcttctaata acctggaatg 6240  
ctgttttccc ggggatcgca gtgggtgagta accatgcac atcaggagta cggataaaat 6300  
gcttgatggt cggaagaggc ataaattccg tcagccagtt tagtctgacc atctcatctg 6360  
taacatcatt ggcaacgcta cctttgccat gtttcagaaa caactctggc gcatcgggct 6420  
tcccatacaa tcgatagatt gtgcacactg attgcccagc attatcgca gccattttat 6480  
acccatataa atcagcatcc atgttggaat ttaatcgcg cctcgagcaa gacgtttccc 6540  
gttgaatatg gtcataaca ccccttgat tactgtttat gtaagcagac agttttattg 6600  
ttcatgatga tatattttta tcttggtgaa tgtaacatca gagattttga gacacaacgt 6660  
ggcttttccc cccccccat tattgaagca tttatcaggg ttattgtctc atgagcggat 6720  
acatatattga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 6780  
aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 6840  
gtatcacgag gccctttcgt c 6861

&lt;210&gt; 87

&lt;211&gt; 7552

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 87

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgttaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggett gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcaatcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt gaaggagaag taccagcacc tgtggcgctg gggctggcgc tggggcacca 1980
tgctgctggg catgctgatg atctgcagcg ccaccgagaa gctgtgggtg accgtgtact 2040
acggcgtgcc cgtgtggcgc gaggccacca ccacctgtt ctgcgccagc gacgccaagg 2100
cctacgacac cgaggtgcac aacgtgtggg ccaccacgc ctgctgccc accgaccca 2160
acccccagga ggtggtgctg ggcaacgtga ccgagaactt caacatgtgg aagaacaaca 2220
tggtggacca gatgcacgag gacatcatca gcctgtggga cgagagcctg aagccctgcg 2280
tgaagctgac cccctgtgc gtgacctga actgcaccaa cctgaacatc accaagaaca 2340
ccaccaacct gaccagcagc agctggggca tgatggagga gggcgagatc aagaactgca 2400
gcttctacat caccaccagc atccgcaaca aggtgaagaa ggagtacgcc ctgttcaacc 2460
gcctggacgt ggtgcccgtg aagaacacca gcaacaccaa gtaccgcctg atcagctgca 2520
acaccagcgt gatcaccag gcctgcccc aagtgagctt ccagccatc cccatccact 2580
actgctgcc cgccggttc gccatcctga agtgcaacaa caagaccttc aacggcagcg 2640
gcccctgcac caacgtgagc accgtgcagt gcaccacgg catccgcccc gtggtgagca 2700
cccagctgct gctgaacggc agcctggccg aggaggacat cgtgatccgc agcgaggact 2760
tcaccgacaa cgtgaagacc atcatcgtgc agctgaacga gagcgtgggt atcaactgca 2820
cccgcccaa caacaacacc cgcgagcgc tgagcatcgg ccccgccgc gccttctacg 2880
cccgccgcaa catcatcgcc gacatccgc agggccactg caacatcagc cgcgccaagt 2940
ggaacaacac cctgcagcag atcgtgatca agctgcgcga gaagtccgc aacaagacca 3000
tcgccttcaa ccagagcagc ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg 3060

```



```

gcggcgagtt cttctactgc aacaccgccc agctgttcaa cagcacctgg aacgtggccg 3120
gcggcaccac cggcaccgag ggcaacgaca tcatcacccct gcagtgccgc atcaagcaga 3180
tcatcaacat gtggcagaag gtgggcaagg ccatgtacgc cccccccatc accggccaga 3240
tccgctgcag cagcaacatc accggcctgc tgcctgacccg cgacggcggc aacagcaccg 3300
agaccgagac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac tggcgagcgc 3360
agctgtacaa gtacaagggt gtgcgcacgc agcccatcgg cgtggccccc acccgcgcca 3420
agcgccgcac cgtgcagcgc gagaagcgcg ccgtgggcct cggcgccgtg ttcctgggct 3480
tcctggggcg cgccggcagc accatggggcg ccgccagcgt gacctgacc gtgcaggccc 3540
gcctgtctgt gagcggcacg gtgcagcagc agaacaacct gctgcgcgcc atcgaggccc 3600
agcagaacat gctgcgcctg accgtgtggg gcatcaagca gctgcaggcc cgcgtgctgg 3660
ccctggagcg ctacctgcgc gaccagcagc tgatgggcct ctggggctgc agcggaagc 3720
ttatctgcac caccagcgtg ccctggaacg tgagctggag caacaagagc gtggacgaca 3780
tctggaacaa catgacctgg atggagtggg agcgcgagat cgacaactac accgactaca 3840
tctacgacct gctggagaag agccagaccc agcaggagaa gaacgagaag gagctgctgg 3900
agctggacaa gtggggcagc ctgtggaact ggttcgacat caccaactgg ctgtggtaca 3960
tccgctgtt catcatgatc gtggggcgcc tgatcggcct gcgcctcgtg ttcgcccgtg 4020
tgagcatcgt gaaccgcgtg cgccagggtt acagccccct gagcttccag accctgctgc 4080
ccgccagccg cgcccccgac cgccccgagg gcaccgagga ggaggcggc gagcgcgacc 4140
gcgaccgcag cgcccccagc gtgaacggca gcctggccct gatctgggac gacctgcgca 4200
gcctgtgcct gttcagctac caccgcctgc gcgacctgct gctgatcgtg acccgcatcg 4260
tggagctgct gggccgcccgc ggctgggagg ccctgaagta ctggtggaac ctgctgcagt 4320
actggagcca ggagctgaag aacagcgccg tgagcctgct gcagtacggc tggagctact 4380
tccacgaggg cgtgcaggcc gtgtggcgca gcgccaccga gacctggcc ggcgccctgg 4440
gcgacctgtg ggagaccctg cgccgcggcg gccgctggat cttggccatc ccccgccgca 4500
tccgccaggg cctggagctg accctgctgt gaacacgtgg gatccagatc tgctgtgcct 4560
tctagtctgc agccatctgt tgtttgcccc tcccccgctg cttccttgac cctggaagg 4620
gccactccca ctgtccttct ctaataaaat gaggaaattg catcgcatct tctgagtagg 4680
tgtcattcta ttttgggggg tggggtgggg caggacagca agggggagga ttgggaagac 4740
aatagcaggg atgctgggga tgcggtgggc tctatgggta cccaggtgct gaagaattga 4800
cccggttctt cctgggcccag aaagaagcag gcacatcccc ttctctgtga cacacctgt 4860
ccacgccccct ggttcttagt tccagcccca ctcataggac actcatagct caggagggct 4920
ccgccttcaa tcccaccgc taaagtactt ggagcggtct ctccctccct catcagccca 4980
ccaaacccaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagt 5040
cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 5100
aaggccatca tggccttaat ctccgccttc ctgcctcact gactcgctgc gctcggctgt 5160
tcggctgcgg cgagcggtat cagctcactc aaaggcggtg atacggttat ccacagaatc 5220
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 5280
aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 5340
tcgacgctca agtcagaggt ggcgaaacc gacaggacta taaagatacc aggcgtttcc 5400
ccctggaagc tccctcgtgc gctctcctgt tccgacctg ccgcttaccg gatacctgtc 5460
cgcttttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 5520
ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga 5580
ccgctgcgcc ttatccggtg actatcgtct tgagtccaac ccggtaaagc acgacttatc 5640
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 5700
agagtctctg aagtgggtgg ctaactacgg ctacactaga agaacagtat ttggtatctg 5760
cgctctgctg aagccagtta ccttcggaaa aagagtgggt agctcttgat ccggcaaaca 5820
aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 5880
aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa 5940
ctcacgttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt 6000
aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaacct ggtctgacag 6060
ttaccaatgc ttaatcagtg aggcacctat ctacgcgac tgtctatttc gttcatccat 6120
agttgcctga ctcggggggg gggggcgctg aggtctgcct cgtgaagaag gtgttgctga 6180
ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc cacggttgat 6240
gagagctttg ttgtagggtg accagttggg gattttgaac ttttgctttg ccacggaacg 6300
gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag ttcgatttat 6360

```

```

tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct gccagtgtta caaccaatta 6420
accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt attcatatca 6480
ggattatcaa taccatattt ttgaaaaagc cgtttctgtg atgaaggaga aaactcaccg 6540
aggcagttcc ataggatggc aagatcctgg tatcgggtctg cgattccgac tcgtccaaca 6600
tcaatacaac ctattaattt cccctcgtca aaaataaggc tatcaagtga gaaatcacca 6660
tgagtgcga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt ccagacttgt 6720
tcaacaggcc agccattacg ctcgatcatca aaatcactcg catcaacca accgttattc 6780
attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg acaattacaa 6840
acaggaatcg aatgcaaccg gcgcaggaac actgccagcg catcaacaat attttcacct 6900
gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc agtgggtgagt 6960
aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg cataaattcc 7020
gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct acctttgcca 7080
tgtttcagaa acaactctgg cgcacgggc ttcccataca atcgatagat tgtcgcacct 7140
gattgcccga cattatcgcg agccatttta taccatata aatcagcatc catglttgaa 7200
tttaatcgcg gcctcgagca agacgtttcc cgttgaatat ggctcataac accccttgta 7260
ttactgttta tgtaagcaga cagttttatt gttcatgatg atatatTTTT atcttggtgca 7320
atgtaacatc agagattttg agacacaacg tggctttccc cccccccca ttattgaagc 7380
atztatcagg gttattgtct catgagcgga tacatatTTg aatgtattta gaaaaataaa 7440
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 7500
attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tc 7552

```

&lt;210&gt; 88

&lt;211&gt; 6774

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 88

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780
cgtcaatggg agtttgTTTT ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccga ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgcctg ccccggtgcc 1380

```

```

cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggte ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcact ggaagcagtg gtggatctgg ggcacccctg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccggt gtggaaggag gccaaagacca ccctgttctg cgcacgcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgctgta 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catggggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggtg cagccctgt tctaccgctt ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgcctca 2520
cccaggcctg cccaagggtg agcttcgacc ccaccccat ccactactgc gccccgcg 2580
gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc cagggacca agcccggtgt gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcaccgc cccaacaaca 2820
acaccgcaa gagcatccgc atcggccccg gccaggcctt ctacgccacc ggcgagatca 2880
tcggcgacat ccgccaggcc cactgcaaca tcagccgcac cgcctggaac aagaccctgc 2940
aggaggtggg caagaagctg gccgagcact tccccaaaaa ggccatcaag ttcgccaagc 3000
acagcggcgg cgacctggag atcaccaccc acagcttcaa ctgccgcggc gagttcttct 3060
actgcaacac cagcagcctg ttcaacagca cctacacccc caacagcacc gagaacatca 3120
ccggcaccga gaacagcatc atcaccatcc cctgccgcac caagcagatc atcaacatgt 3180
ggcagggcgt gggccgcgcc atgtacgcc ccccatcga gggcatcctg acctgccgca 3240
gcaacatcac cggcctgctg ctgacccgcg acggcggcac cggcatgcac gacaccgaga 3300
tcttccgctt gggcggcggc gacatgcgcg acaactggcg cagcgagctg tacaagtaca 3360
agggtggtgaa gatcgagccc ctgggcgtgg cccccaccaa ggccaagctt accgtgcagg 3420
cccgccagct gctgagcggc atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg 3480
cccagcagca cctgctgcag ctgaccgtgt ggggcaccaa gcagctgcag gcccgcaccc 3540
tggccgtgga gcgctacctg aaggaccagc agctggagca gatctggaac cacaccacct 3600
ggatggagtg ggaccgcgag atcaacaact acaccagcct gatccacagc ctgatcgagg 3660
agagccagaa ccagcaggag aagaacgagc aggagctgct ggagctggac aagtgggcca 3720
gcctgtggaa ctggttcaac atcaccaact ggctgtggtg aggatccaga tctgctgtgc 3780
cttctagtgt ccagccatct gttgtttgcc cctccccgt gccttccttg accctggaag 3840
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcac tgtctgagta 3900
ggtgtcatte tattctgggg ggtggggtgg ggcaggacag caaggggggag gattgggaag 3960
acaatagcag gcatgctggg gatgcggtgg gctctatggg taccaggtg ctgaagaatt 4020
gacccggttc ctctggggcc agaaagaagc aggcacatcc ccttctctgt gacacaccct 4080
gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag ctgaggagg 4140
ctccgccttc aatccacccc gctaaagtac ttggagcggc ctctccctcc ctcatcagcc 4200
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4260
tgcagagggg gagaaaatgc ctccaacatg tgaggaagta atgagagaaa tcatagaatt 4320
ttaaggccat catggcctta atcttccgt tctcgtccta ctgactcgtc gcgctcggtc 4380
gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg taatacgggt atccacagaa 4440
tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 4500
aaaaaggccg cgttgcctggc gtttttccat aggtccgccc cccctgacga gcatcacaaa 4560
aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 4620
ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 4680

```



```

tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 4740
agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 4800
gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtatg acacgactta 4860
tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcgggtgct 4920
acagagttct tgaagtgggtg gcctaactac ggctacacta gaagaacagt atttggtatc 4980
tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 5040
caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac gcgcagaaaa 5100
aaaggatctc aagaagatcc ttgatcttt tctacggggt ctgacgctca gtggaacgaa 5160
aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 5220
ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 5280
agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc 5340
atagttgcct gactcggggg ggggggggcg tgaggtctgc ctctgaaga aggtgttgct 5400
gactcatacc aggcctgaat cgcctcatca tccagccaga aagtgaggga gccacgggtg 5460
atgagagctt tgttgtaggt ggaccagttg gtgattttga acttttgctt tgccacggaa 5520
cggctctcgt tgtcgggaag atgcgtgac tgatccttca actcagcaaa agttcgattt 5580
attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtg tacaaccaat 5640
taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatat 5700
caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac 5760
cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtccaa 5820
catcaatata acctattaat tccccctcgt caaaaataag gttatcaagt gagaaatcac 5880
catgagtgc gactgaatcc ggtgagaatg gcaaaagctt atgcatttct ttccagactt 5940
gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcataacc aaaccgttat 6000
tcattcgtga ttgcgcctga gcgagacgaa atacgcgac gctgttaaaa ggacaattac 6060
aaacaggaat cgaatgcaac cggcgcagga acactgccag cgcateaca atattttcac 6120
ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc gcagtgggtg 6180
gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcgggaaga ggcataaatt 6240
ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc 6300
catgtttcag aaacaactct ggcgcacggt gcttcccata caatcgatag attgtcgac 6360
ctgattgcc gacattatcg cgagccatt tatacccata taaatcagca tccatggttg 6420
aatttaatcg cggcctcgag caagacgttt cccgttgaat atggctcata acacccttg 6480
tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatatt ttatcttggtg 6540
caatgtaaca tcagagattt tgagacacaa cgtggctttt ccccccccc cattattgaa 6600
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 6660
aacaatatag ggttcgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca 6720
ttattatcat gacattaacc tataaaaata ggcgtatcac gaggccctt cgtc 6774

```

&lt;210&gt; 89

&lt;211&gt; 6846

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 89

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480

```

catagtaacg	ccaatagggg	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaa	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgcc	ggcattatgc	ccagtacatg	accttatggg	actttccctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatbga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccccttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttaciaaatt	cacatataca	acaacgcctg	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttagggac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcgggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatacgcg	ccgctctaga	gatatacgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgatc	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaagacca	ccctgttctg	cgcagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtgggcca	cccacgcctg	cgtgcccacc	gacccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	caacatgacc	accgaggtgc	2400
gcgacaagcg	cgagaagggtg	cacgcccctgt	tctaccgcct	ggacatcgtg	cccctgaaga	2460
acgagagcag	caacaccagc	ggcgactacc	gcctgatcaa	ctgcaacacc	agcgccatca	2520
cccaggcctg	ccccaaagggtg	agcttcgacc	ccatccccat	ccactactgc	gccccgcgcg	2580
gctacgccat	cctgaagtgc	aacaacaaga	ccttcaacgg	caccggcccc	tgcaacaacg	2640
tgagcaccat	ccagtgcacc	cacggcacca	agcccggtgt	gagcaccag	ctgctgctga	2700
acggcagcct	ggccgaggag	gagatcatca	tccgcagcaa	gaacctgacc	gacaacgtga	2760
agaccatcat	cgtgcacctg	aacgagagcg	tggagatcaa	ctgcaccgcg	ccaacaaca	2820
acacccgcaa	gagcatccgc	atcggccccg	gccaggcctt	ctacgccacc	ggcgagatca	2880
tcggcgacat	ccgccaggcc	cactgcaaca	tcagccgcac	cgcctggaac	aagaccctgc	2940
aggaggtggg	caagaagctg	gccgagcact	tccccaaaca	ggccatcaag	ttcgccaagc	3000
acagcggcgg	cgacctggag	atcaccaccc	acagcttcaa	ctgcgcgggc	gagttcttct	3060
actgcaacac	cagcagcctg	ttcaacagca	cctacacccc	caacagcacc	gagaacatca	3120
ccggcaccga	gaacagcatc	atcaccatcc	cctgccgcct	caagcagatc	atcaacatgt	3180
ggcagggcgt	gggcgcgcgc	atgtacgccc	ccccatcga	gggcatcctg	acctgccgca	3240
gcaacatcac	cggcctgctg	ctgacccgcg	acggcggcac	cggcatgcac	gacaccgaga	3300
tcttccgcct	gggcggcggc	gacatgcgcg	acaactggcg	cagcgagctg	tacaagtaca	3360
aggtgggtgaa	gatcgagccc	ctgggcgtgg	ccccaccaa	ggccaagctt	accgtgcagg	3420
cccgccagct	gctgagcggc	atcgtgcagc	agcagaacaa	cctgctgcgc	gccatcgagg	3480
cccagcagca	cctgctgcag	ctgaccgtgt	ggggcatcaa	gcagctgcag	gcccgcctcc	3540
tggccgtgga	gcgctacctg	aaggaccagc	agctggagca	gatctggaac	cacaccacct	3600
ggatggagtg	ggaccgcgag	atcaacaact	acaccagcct	gatccacagc	ctgatcgagg	3660
agagccagaa	ccagcaggag	agaacgagc	aggagctgct	ggagctggac	aagtgggcca	3720
gcctgtggaa	ctgggttcaac	atcaccaact	ggctgtggta	catcaagctg	ttcatcatga	3780

tcgtgggagg	cctgggtggg	ctgcgcacgc	tgttcgcccgt	gctgagcatc	tgaggatcca	3840
gatctgctgt	gccttctagt	tgccagccat	ctggtggttg	ccccccccc	gtgccttcc	3900
tgaccctgga	aggtgccact	cccactgtcc	tttcctaata	aaatgaggaa	attgcatcgc	3960
attgtctgag	taggtgtcat	tctattctgg	ggggtgggg	ggggcaggac	agcaagggg	4020
aggattggga	agacaatagc	aggcatgctg	gggatgcggt	gggctctatg	ggtacccagg	4080
tgctgaagaa	ttgaccgggt	tcctcctggg	ccagaaagaa	gcaggcacat	ccccctctct	4140
gtgacacacc	ctgtccacgc	ccctggttct	tagttccagc	cccactcata	ggacactcat	4200
agctcaggag	ggctccgcct	tcaatcccac	ccgctaaagt	acttggagcg	gtctctccct	4260
ccctcatcag	cccaccaaac	caaacctagc	ctccaagagt	gggaagaaat	taaagcaaga	4320
taggctatta	agtgcagagg	gagagaaaat	gcctccaaca	tgtgaggaag	taatgagaga	4380
aatcatagaa	ttttaaggcc	atcatggcct	taatcttccg	cttcctcgct	cactgactcg	4440
ctgcgctcgg	tcgttcgggt	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	4500
ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	4560
gccaggaacc	gtaaaaaggc	cgcggttgctg	gcgtttttcc	ataggctccg	ccccctgac	4620
gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	4680
taccaggcgt	ttccccctgg	aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	4740
accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtag	cgctttctca	tagctcacgc	4800
tgtaggatc	tcagttcggt	gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	4860
cccgttcagc	ccgaccgctg	cgcttatcc	ggtaactatc	gtcttgagtc	caacccggta	4920
agacacgact	tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	4980
gtaggcgggt	ctacagaggt	cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	5040
gtatttggtg	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	5100
tgatccggca	aacaaaccac	cgctggtagc	gggtggtttt	ttgtttgcaa	gcagcagatt	5160
acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	5220
cagtggaaac	aaaactcacg	ttaagggtatt	ttggtcatga	gattatcaaa	aaggatcttc	5280
acctagatcc	ttttaaatga	aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	5340
acttggtctg	acagttacca	atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	5400
tttcggtcat	ccatagttgc	ctgactcggg	gggggggggc	gctgagggtc	gcctcgtgaa	5460
gaagggtggt	ctgactcata	ccaggcctga	atcgccccat	catccagcca	gaaagtgagg	5520
gagccacggg	tgatgagagc	tttggtgtag	gtggaccagt	tgggtgattt	gaacttttgc	5580
tttgccacgg	aacgggtctg	gttggtcgga	agatgcgtga	tctgatcctt	caactcagca	5640
aaagttcgat	ttattcaaca	aagccgcggt	cccgtaagt	cagcgtaatg	ctctgccagt	5700
gttacaacca	attaaccaat	tctgattaga	aaaactcatc	gagcatcaaa	tgaaactgca	5760
atattattcat	atcaggatta	tcaataccat	atttttgaaa	aagccgtttc	tgtaatgaag	5820
gagaaaactc	accgaggcag	ttccatagga	tggcaagatc	ctgggtatcg	tctgcgattc	5880
cgactcgtcc	aacatcaata	caacctatta	atttccctc	gtcaaaaata	aggttatcaa	5940
gtgagaaatc	accatgagtg	acgactgaat	ccggtgagaa	tggcaaaagc	ttatgcattt	6000
ctttccagac	ttgttcaaca	ggccagccat	tacgctcgtc	atcaaaatca	ctcgcatcaa	6060
ccaaaccggt	attcattcgt	gattgcgctt	gagcgagacg	aaatacgcg	tcgctgttaa	6120
aaggacaatt	acaaacagga	atcgaatgca	accggcgagc	gaacactgcc	agcgcatcaa	6180
caatattttc	acctgaatca	ggatattctt	ctaatacctg	gaatgctgtt	ttcccgggga	6240
tcgcagtggt	gagtaaccat	gcatacatcag	gagtacggat	aaaatgcttg	atggtcggaa	6300
gaggcataaa	ttccgtcagc	cagtttagtc	tgaccatctc	atctgtaaca	tcattggcaa	6360
cgctaccttt	gccatgtttc	agaaacaact	ctggcgcatc	gggcttccca	tacaatcgat	6420
agattgtcgc	acctgattgc	ccgacattat	cgcgagccca	tttataacca	tataaatcag	6480
catccatggt	ggaatttaat	cgcgccctcg	agcaagacgt	ttcccgttga	atatggctca	6540
taacacccct	tgtattactg	tttatgtaag	cagacagttt	tattgttcat	gatgatatat	6600
ttttatcttg	tgcaatgtaa	catcagagat	tttgagacac	aacgtggctt	ttccccccc	6660
cccattattg	aagcatttat	cagggttatt	gtctcatgag	cggatacata	tttgaatgta	6720
tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	6780
tctaagaaac	cattattatc	atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	6840
ttcgtc						6846

&lt;210&gt; 90

&lt;211&gt; 6853

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 90

```
tgcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatglt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatltt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcaccttgg 1980
gcttctggat ggtgatgata tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cgcagcgcac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggtg cacgccctgt tctaccgcct ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520
cccaggcctg ccccaagggtg agcttcgacc ccattcccat ccactactgc gccccgcg 2580
gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc cagggcacca agcccggtgt gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
```



agaccatcat	cgtgcacctg	aacgagagcg	tggagatcaa	ctgcacccgc	cccaacaaca	2820
acacccgcaa	gagcatccgc	atcggccccg	gccaggcctt	ctacgccacc	ggcgagatca	2880
tcggcgacat	ccgccaggcc	cactgcaaca	tcagccgcac	cgcctggaac	aagaccctgc	2940
aggaggtggg	caagaagctg	gccgagcact	tccccaacaa	ggccatcaag	ttcgccaagc	3000
acagcggcgg	cgacctggag	atcaccaccc	acagcttcaa	ctgccgcggc	gagttcttct	3060
actgcaacac	cagcagcctg	ttcaacagca	cctacacccc	caacagcacc	gagaacatca	3120
ccggcacccga	gaacagcate	atcaccatcc	cctgccgcct	caagcagatc	atcaacatgt	3180
ggcagggcgt	gggccgcgcc	atgtacgccc	cccccatcga	gggcatacctg	acctgccgca	3240
gcaacatcac	cggcctgctg	ctgacccgcg	acggcggcac	cggcatgcac	gacaccgaga	3300
tcttccgccc	cgagggcggc	gacatgcgcg	acaactggcg	cagcgagctg	tacaagtaca	3360
aggtggtgga	gatcaagccc	ctgggcatcg	ccccaccaaa	ggccaagctt	accgtgcagg	3420
tgcgccagct	gctgagcggc	atcgtgcagc	agcagagcaa	cctgctgcgc	gccatcgagg	3480
cccagcagca	catgctgcag	ctgaccgtgt	ggggcatcaa	gcagctgcag	acccgcgtgc	3540
tggccatcga	gcgctacctg	cgcgaccagc	agctcgagga	catctggaac	aacatgacct	3600
ggatgcagtg	ggaccgcgag	atcagcaact	acaccaacac	catctaccgc	ctgctggagg	3660
acagccagaa	ccagcaggag	aagaacgagc	aggacctgct	ggccctggac	aagtggcaga	3720
acctgtggac	ctgggttcggc	atcaccaact	ggctgtggta	catcaagatc	ttcatcaaga	3780
tcgtgggcgg	cctgatcggc	ctgcgcatca	tcttcgccgt	gctgagcatc	tgaacacgtg	3840
ggatccagat	ctgctgtgcc	ttctagtctg	cagccatctg	ttgtttgccc	ctcccccgctg	3900
ccttccttga	ccctggaagg	tgccactccc	actgtccttt	cctaataaaa	tgaggaaatt	3960
gcatecgcat	gtctgagtag	gtgtcattct	attctggggg	gtgggggtggg	gcaggacagc	4020
aagggggagg	attgggaaga	caatagcagg	catgctgggg	atgcggtggg	ctctatgggt	4080
accaggtgc	tgaagaattg	acccggttcc	tcttgggcca	gaaagaagca	ggcacatccc	4140
cttctctgtg	acacaccctg	tccacgcccc	tggttcttag	ttccagcccc	actcatagga	4200
cactcatagc	tcaggagggg	tccgccttca	atcccacccg	ctaaagtact	tggagcggtc	4260
tctccctccc	tcateagccc	accaaaccaa	acctagcctc	caagagtggg	aagaaattaa	4320
agcaagatag	gctattaagt	gcagagggag	agaaaatgcc	tccaacatgt	gaggaagtaa	4380
tgagagaaat	catagaattt	taaggccatc	atggccttaa	tcttccgctt	cctcgcctcac	4440
tgactcgctg	cgctcggtcg	ttcggtcgcg	gcgagcggta	tcagctcact	caaaggcggc	4500
aatacgggta	tccacagaat	caggggataa	cgcaggaaaag	aacatgtgag	caaaaggcca	4560
gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	4620
ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	4680
ataaagatac	caggcggtttc	ccccctggaag	ctccctcgctg	cgctctcctg	ttccgacctt	4740
gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	4800
ctcacgctgt	aggtatctca	gttcgggtgta	ggctcgttcgc	tccaagctgg	gctgtgtgca	4860
cgaaccccc	gttcagcccc	accgctgcgc	cttatccggc	aactatcgtc	ttgagtccaa	4920
cccggtaaga	cacgacttat	cgcactggc	agcagccact	ggtaacagga	ttagcagagc	4980
gaggtatgta	ggcgggtgta	cagagtctct	gaagtgggtg	cctaactacg	gctacactag	5040
aagaacagta	tttgggtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttag	5100
tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggc	ggtttttttg	tttgcaagca	5160
gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	5220
tgacgctcag	tggaaacgaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	5280
gatcttcacc	tagatccttt	taaattaaaa	atgaagtttt	aatcaatct	aaagtatata	5340
tgagtaaaact	tggctctgaca	gttaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	5400
ctgtctattt	cgttcatcca	tagttgcctg	actcgggggg	ggggggcgct	gaggtctgcc	5460
tcgtgaagaa	ggtgttgctg	actcatacca	ggcctgaatc	gccccatcat	ccagccagaa	5520
agtgaaggag	ccacgggtga	tgagagcttt	gttgtaggtg	gaccagttag	tgattttgaa	5580
cttttgcttt	gccacggaac	ggtctgcgtt	gtcgggaaga	tgcgtgatct	gaccttcaa	5640
ctcagcaaaa	gttcgattta	ttcaacaaag	ccgccgtccc	gtcaagtcag	cgtaatgctc	5700
tgccagtgtt	acaaccaatt	aaccaattct	gattagaaaa	actcatcgag	catcaaatga	5760
aactgcaatt	tattcatatc	aggattatca	ataccatatt	tttgaaaaag	ccgtttctgt	5820
aatgaaggag	aaaactcacc	gaggcagttc	cataggatgg	caagatcctg	gtatcgggtct	5880
gcgattccga	ctcgtccaac	atcaatacaa	cctattaatt	tccctcgtc	aaaaataagg	5940
ttatcaagtg	agaaatcacc	atgagtgcag	actgaatccg	gtgagaatgg	caaaagctta	6000
tgcatttctt	tccagacttg	ttcaacaggc	cagccattac	gctcgtcatc	aaaatcactc	6060



```

gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg 6120
ctgtttaaag gacaattaca aacaggaatc gaatgcaacc ggcgaggaa cactgccagc 6180
gcatcaacaa ttttttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc 6240
ccggggatcg cagtgggtgag taacctatgca tcatcaggag tacggataaa atgcttgatg 6300
gtcgggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca 6360
ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcacgagg cttcccatac 6420
aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccatth ataccatat 6480
aaatcagcat ccatgttgga atttaatcgc ggccctcagc aagacgtttc ccgttgaata 6540
tggtcctata cacccttgt attactgttt atgtaagcag acagttttat tgttcatgat 6600
gatataatct tatcttgtgc aatgtaacat cagagattht gagacacaac gtggctttcc 6660
ccccccccc attattgaag catttatcag ggtrattgtc tcatgagcgg atacataatth 6720
gaatgtatth agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc 6780
cctgacgtct aagaaacctt tattatcatg acattaacct ataaaaatag gcgtatcacg 6840
agccctttc gtc 6853

```

&lt;210&gt; 91

&lt;211&gt; 6761

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 91

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aacgcccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaggt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcatatgac ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatth ccaagtctcc accccattga 780
cgtcaatggg agtttgthtt ggacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg ggccctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggctcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcaac gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	gcgcggcatc	cagaccagct	ggcagaacct	gtggcgctgg	ggcaccatga	1980
tcctgggcat	gctggtgatc	tacagcgccg	ccgagaacct	gtgggtggcc	gtgtactacg	2040
gcgtgcccg	gtggaaggac	gccgagacca	ccctgttctg	cgccagcgac	gccaaggcct	2100
acgacaccga	ggtgcacaac	gtgtgggaga	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagg	ggacctggag	aacgtgaccg	aggacttcaa	catgtggcgc	aacaacatgg	2220
tggagcagat	gcacaccgac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctggact	gcaacgccac	cgccagcaac	gtgaccaacg	2340
agatgcgcaa	ctgcagcttc	aacatcacca	ccgagctgaa	ggacaagaag	cagcaggtgt	2400
acagcctgtt	ctacaagctg	gacgtgggtg	agatcaacga	gaagaacgag	accgacaagt	2460
accgcctgat	caactgcaac	accagcgcca	tcacccaggc	ctgccccaa	gtgagcttcg	2520
agcccatccc	catccactac	tgcgcccccg	ccggcttcgc	catcctgaag	tgcaaggaca	2580
ccgagttcaa	cggcaccggc	ccctgcaaga	acgtgagcac	cgtgcagtgc	accacgggca	2640
tccgccccgt	gatcagcacc	cagctgctgc	tgaacggcag	cctggccgag	gagggcatcc	2700
agatccgcag	cgagaacatc	accaacaacg	ccaagaccat	catcgtgcag	ctggacaagg	2760
ccgtgaagat	caactgcacc	cgcaccaaca	acaacacccg	caagggcgtg	cgcatcggcc	2820
ccggccaggc	cttctacgcc	accggcggca	tcatcggcga	catccgccag	gcccactgcc	2880
acgtgagccg	cgccaagtgg	aacgacaccc	tgcgcggcgt	ggccaagaag	ctgcgcgagc	2940
acttcaagaa	caagaccatc	atcttcgaga	agagcagcgg	cggcgacatc	gagatcacca	3000
cccacagctt	catctgcggc	ggcgagttct	tctactgcaa	caccagcggc	ctgttcaaca	3060
gcacctggga	gagcaacagc	accgagagca	acaacaccac	cagcaacgac	accatcaccc	3120
tgacctgccg	catcaagcag	atcatcaaca	tgtggcagaa	ggtgggcccag	gccatgtacc	3180
cccccccat	ccagggcgtg	atccgctgcg	agagcaacat	caccggcctg	ctgctgaccc	3240
gcgacggcgg	caacaacagc	accaacgaga	tcttccgccc	cggcgggcggc	aacatgcgcg	3300
acaactggcg	cagcgagctg	tacaagtaca	aggtgggtgaa	gatcgagccc	ctgggcgtgg	3360
ccccagccg	cgccaagctt	accgcccagg	cccggccagt	gctgagcggc	atcgtgcagc	3420
agcagagcaa	cctgctgcgc	gccatcgagg	cccagcagca	catgctgaag	ctgaccgtgt	3480
ggggcatcaa	gcagctgcag	gcccgcgtgc	tggccgtgga	gcgctacctg	aaggaccagc	3540
agctcgagat	ctgggacaac	atgacctggc	tgcagtggga	caaggagatc	agcaactaca	3600
cccagatcat	ctacaacctg	atcgaggaga	gccagaacca	gcaggagaag	aacgagcagg	3660
acctgctggc	cctggacaag	tgggcccagcc	tgtggaactg	gttcgacatc	agccgctggc	3720
tgtggtgaac	acgtgggatc	cagatctgct	gtgccttcta	gttgccagcc	atctgttgtt	3780
tgccctccc	ccgtgccttc	cttgacctg	gaaggtgcca	ctcccactgt	cctttcctaa	3840
taaaatgagg	aaattgcate	gcattgtctg	agtaggtgtc	attctattct	ggggggtggg	3900
gtggggcagg	acagcaaggg	ggaggattgg	gaagacaata	gcaggcatgc	tggggatgcg	3960
gtgggctcta	tgggtaccca	ggtgctgaag	aattgacccg	gttcctcctg	ggccagaaag	4020
aagcaggcac	atcccttct	ctgtgacaca	ccctgtccac	gcccctgggt	cttagttcca	4080
gccccactca	taggacactc	atagctcagg	agggctccgc	cttcaatccc	accgctaaa	4140
gtacttgagg	cggctctctc	ctccctcatc	agcccaccaa	accaaacctc	gcctccaaga	4200
gtgggaagaa	attaaagcaa	gataggctat	taagtgcaga	gggagagaaa	atgcctccaa	4260
catgtgagga	agtaatgaga	gaaatcatag	aattttaagg	ccatgattta	aggccatcat	4320
ggccttaate	ttccgcttcc	tcgctcactg	actcgctgcg	ctcggctcgt	cggtgcggc	4380
gagcgggtatc	agctcactca	aaggcggtaa	tacgggttatc	cacagaatca	ggggataacg	4440
caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	4500
tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	4560
gtcagagggtg	gcgaaaccgc	acaggactat	aaagatacca	ggcgtttccc	cctggaagct	4620
ccctcgctgcg	ctctcctggt	ccgaacctgc	cgttaccgg	atacctgtcc	gcctttctcc	4680
cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	4740
tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccgt	tcagcccgac	cgctgcgcct	4800
tatccggtaa	ctatcgctct	gagtccaacc	cggtaagaca	cgacttatcg	ccactggcag	4860
cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	cggtgctaca	gagttcttga	4920
agtgggtggcc	taactacggc	tacactagaa	gaacagtatt	tggtatctgc	gctctgctga	4980
agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaa	accacgcgtg	5040

```

gtagcgggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag 5100
aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaac tcacgttaag 5160
ggattttggg catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat 5220
gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct 5280
taatcagtga ggcacctatc tcagcgatct gtctatttctg ttcattccata gttgcctgac 5340
tcgggggggg ggggcgctga ggtctgcctc gtgaagaagg tgttgctgac tcataccagg 5400
cctgaatcgc cccatcatcc agccagaaaag tgaggaggcc acggttgatg agagctttgt 5460
tgtagggtga ccagtttggtg attttgaact tttgctttgc cacggaacgg tctgcgttgt 5520
cgggaagatg cgtgatctga tccttcaact cagcaaaagt tcgatttatt caacaaagcc 5580
gccgtcccg caagtcagcg taatgctctg ccagtgttac aaccaattaa ccaattctga 5640
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat 5700
accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 5760
taggatggca agatcctggg atcgggtctgc gattccgact cgtccaacat caatacaacc 5820
tattaatttc cctcgtcaa aaataagggt atcaagttag aaatcaccat gagtgcgac 5880
tgaatccggt gagaatggca aaagcttatg catttcttcc cagacttggt caacaggcca 5940
gccattacgc tcgtcatcaa aatcactcgc atcaacccaa ccgttattca ttcgtgattg 6000
cgcctgagcg agacgaaata cgcgatcgtg gttaaaagga caattacaaa caggaatcga 6060
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata 6120
ttcttcta atcctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcac 6180
atcaggagta cggataaaat gcttgatggg cgggaagggc ataaattccg tcagccagtt 6240
tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa 6300
caactctggc gcacggggct tcccatataa tcgatagatt gtgcgacctg attgcccagc 6360
attatcgca gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcg 6420
cctcgagcaa gacgtttccc gttgaatatg gtcataaca ccccttgat tactgtttat 6480
gtaagcagac agttttattg ttcattgatg tatattttta tcttggtgca tgtaacatca 6540
gagattttga gacacaacgt ggctttcccc cccccccat tattgaagca tttatcaggg 6600
ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 6660
tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac 6720
attaacctat aaaaataggc gtatcacgag gccctttcgt c 6761

```

&lt;210&gt; 92

&lt;211&gt; 6836

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 92

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttccctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```



ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgaag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattgggtat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	ttttacaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	atataacata	gcgtgggagc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggaggggcag	gtagtctgag	cagtactcgt	tgtgtccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcccttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatacgcg	ccgctctaga	gatatacgcca	1920
ccatgcgcgt	gcgcggcatc	cagaccagct	ggcagaacct	gtggcgctgg	ggcaccatga	1980
tccctgggcat	gctgggtgatc	tacagcgccg	ccgagaacct	gtgggtggcc	gtgtactacg	2040
gcgtgcccgt	gtggaaggac	gccgagacca	ccctgttctg	cgcagcgac	gccaaggcct	2100
acgacaccga	ggtgcacaac	gtgtgggaga	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggaggt	ggacctggag	aacgtgaccg	aggacttcaa	catgtggcgc	aacaacatgg	2220
tggagcagat	gcacaccgac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctggact	gcaacgccac	cgcagcaac	gtgaccaacg	2340
agatgcgcaa	ctgcagcttc	aacatcacca	ccgagctgaa	ggacaagaag	cagcaggtgt	2400
acagcctgtt	ctacaagctg	gacgtgggtg	agatcaacga	gaagaacgag	accgacaagt	2460
accgcctgat	caactgcaac	accagcgcca	tcacccagge	ctgccccaa	gtgagcttcg	2520
agcccatccc	catccactac	tgcgcccccg	ccggcttcgc	catcctgaag	tgcaggaca	2580
ccgagttcaa	cggcaccggc	ccctgcaaga	acgtgagcac	cgtgcagtgc	accacggcca	2640
tccgccccgt	gatcagcacc	cagctgctgc	tgaacggcag	cctggccgag	gagggcatcc	2700
agatccgcag	cgagaacatc	accaacaacg	ccaagaccat	catcgtgcag	ctggacaagg	2760
ccgtgaagat	caactgcacc	cgcaccaaca	acaacacccg	caagggcgtg	cgcacggccc	2820
ccggccaggc	cttctacgcc	accggcggca	tcacgcggca	catccgccag	gcccactgcc	2880
acgtgagccg	cgcgaagtgg	aacgacaccc	tgcgcggcgt	ggccaagaag	ctgcgcgagc	2940
acttcaagaa	caagaccatc	atcttcgaga	agagcagcgg	cggcgacatc	gagatcacca	3000
cccacagctt	catctgcggc	ggcgagttct	tctactgcaa	caccagcggc	ctgttcaaca	3060
gcacctggga	gagcaacagc	accgagagca	acaacaccac	cagcaacgac	accatcaccc	3120
tgacctgccg	catcaagcag	atcatcaaca	tgtggcagaa	gggtgggccag	gccatgtacc	3180
cccccccat	ccagggcgtg	atccgctgcg	agagcaacat	caccggcctg	ctgctgaccc	3240
gcgacggcgg	caacaacagc	accaacgaga	tcttccgccc	cggcggcggc	aacatgcgcg	3300
acaactggcg	cagcgagctg	tacaagtaca	aggtgggtgaa	gatcgagccc	ctgggcgtgg	3360
ccccagccc	cgcgaagctt	accgcccagg	cccgcagct	gctgagcggc	atcgtgcagc	3420
agcagagcaa	cctgctgcgc	gccatcgagg	cccagcagca	catgctgaag	ctgaccgtgt	3480
ggggcatcaa	gcagctgcag	gcccgcgtgc	tggccgtgga	gcgctacctg	aaggaccagc	3540
agctcgagat	ctgggacaac	atgacctggc	tgcagtggga	caaggagatc	agcaactaca	3600
cccagatcat	ctacaacctg	atcgaggaga	gccagaacca	gcaggagaag	aacgagcagg	3660
acctgctggc	cctggacaag	tgggccagcc	tgtggaactg	gttcgacatc	agccgctggc	3720
tgtggtagat	caagatcttc	atcatgatcg	tgggcggcct	gatcggcctg	cgcacgtgtt	3780
tcgcccgtgt	gagcgtgatc	tgaacacgtg	ggatccagat	ctgctgtgcc	ttctagtgtc	3840
cagccatctg	ttgtttgccc	ctcccccggtg	ccttccttga	ccctggaagg	tgccactccc	3900
actgtccttt	cctaataaaa	tgaggaaatt	gcacgcattt	gtctgagtag	gtgtcattct	3960
attctggggg	gtgggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	4020
catgctgggg	atgcgggtggg	ctctatgggt	accaggtgc	tgaagaattg	accgggttcc	4080
tcctggggcca	gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	tccacgcccc	4140

tggttcttag ttccagcccc actcatagga cactcatage tcaggagggc tccgccttca 4200  
atcccacccg ctaaagtact tggagcggtc tctccctccc tcatcagccc accaaaccaa 4260  
acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag 4320  
agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg 4380  
atttaaggcc atcatggcct taatcttccg ctctctcgct cactgactcg ctgcgctcgg 4440  
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 4500  
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 4560  
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca 4620  
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 4680  
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 4740  
tgtecgctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 4800  
tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc 4860  
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta agacacgact 4920  
tategccact ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcgggtg 4980  
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg 5040  
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 5100  
aacaaccac cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa 5160  
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg 5220  
aaaactcacg ttaagggtt ttggtcatga gattatcaaa aaggatcttc acctagatcc 5280  
ttttaaatga aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 5340  
acagttacca atgcttaatc agtgaggcac ctatctcage gatctgtcta tttcgttcat 5400  
ccatagttgc ctgactcggg gggggggggc gctgaggctt gcctcgtgaa gaagggtgtg 5460  
ctgactcata ccaggcctga atcgcccat catccagcca gaaagtgagg gagccacggg 5520  
tgatgagagc tttgtttagt gtggaccagt tggtagtttt gaacttttgc tttgccacgg 5580  
aacggtctgc gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat 5640  
ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca 5700  
attaaccaat tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat 5760  
atcaggatta tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc 5820  
accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc 5880  
aacatcaata caacctatta atttccctc gtcaaaaata aggttatcaa gtgagaaatc 5940  
accatgagtg acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac 6000  
ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt 6060  
attcattcgt gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt 6120  
acaaacagga atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc 6180  
acctgaatca ggatattctt ctaatacctg gaatgctgtt tccccgggga tcgcagtggg 6240  
gagtaaccat gcatcatcag gactacggat aaaatgcttg atggtcggaa gaggcataaa 6300  
ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt 6360  
gccatgttcc agaaacaact ctggcgcacg gggcttccca tacaatcgat agattgtcgc 6420  
acctgattgc ccgacattat cgcgagccca tttataacca tataaatcag catccatgtt 6480  
ggaatttaat cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccctt 6540  
tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg 6600  
tgcaatgtaa catcagagat tttgagacac aacgtggctt tccccccccc cccattattg 6660  
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6720  
taaacaataa ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6780  
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6836

&lt;210&gt; 93

&lt;211&gt; 6800

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes



&lt;400&gt; 93

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtctt gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtaaatggg agttttgttt ggcacccaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtggtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgcgt gaaggagacc cagatgaact ggcctaacct gtggaagtgg ggcaccccta 1980
tcctgggcct ggtgatcatc tgcagcgcca gcgacaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggcgcgac gccgacacca ccctgttctg cgccagcgac gccaaaggcc 2100
acgagaccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gacccaacc 2160
cccaggaggt ggacctggag aacatcaccc agaacttcaa catgtggaag aacaacatgg 2220
tggagcagat gcaggaggac gtgatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgaact gcatcaacgc caacctgacc aacgtgaaca 2340
acatcatcca cggccccaac atcatcgcca acatcaccca cgagggtgcgc aactgcagct 2400
tcaacatgac caccgagatc cgcgacaaga agcagaaggt gcacgccctg ttctacaagc 2460
tggacctggg gcagatcgag gacaagaaca gcagcgagta ccgctgatc aactgcaaca 2520
ccagcgtgat caagcaggcc tgccccaaga tcagcttcga ccccatcccc atccactact 2580
gcacccccgc cggctacgcc atcctgaagt gcaacgacaa gaacttcaac ggcaccggcc 2640
cctgcaccaa cgtgagcagc gtgcagtga cccacggcat caagcccgtg gtgagcacc 2700
agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc gagaacctga 2760
ccaacaacgc caagaccatc atcgtgcacc tgaacaagag cgtggagatc aactgcacc 2820
gccccagcaa caacaccgc accagacca ccatcgccc cgccaggtg ttctaccgca 2880
ccggcgacat caccggcaac atccgcaagg cctactgcga gatcaacggc accaagtgga 2940
acgaggccct gaagcagggt accaagaagc tgaaggagca cttcaacaac aagaccatca 3000
tcttccagcc ccccgcggtt ggcgacctgg agatcaccat gcaccacttc aactgccgcg 3060
gcgagttctt ctactgcaac accacccgcc tgttcaacaa cacctgcac ggcaacgaga 3120
ccatcggcgg ctgcaacgac accatcatcc tgccctgcaa gatcaagcag atcatcaaca 3180

```

tgtggcaggg	cgccggcctg	gccatgtacg	cccccccat	cagcggccgc	atcaactgcg	3240
tgagcaacat	caccggcatc	ctgctgaccc	gcgacggcgg	cgccaacaac	accagcaacg	3300
agatcttccg	ccccggcggc	ggcaacatca	aggacaactg	gcgcagcgag	ctgtacaagt	3360
acaaggtggt	gcagatcgag	cccctgggca	tcgccccac	ccgcgccaag	cttaccgtgc	3420
aggcccggca	gctggtgagc	ggcatcgtgc	agcagcagag	caacctgctg	cgcgccatcg	3480
aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	caggcccgcg	3540
tgctggccgt	ggagcgctac	ctgaaggacc	agaagttcct	cgaggagatc	tggagcaaca	3600
tgacctggat	cgagtgggag	cgcgagatca	gcaactacac	caaccagatc	tacgagatcc	3660
tgaccgagag	ccagaaccag	caggaccgca	acgagaagga	cctgctggag	ctggacaagt	3720
gggcgagcct	gtggagctgg	ttcgacatca	ccaactggct	gtggtgaaca	cgtgggatcc	3780
agatctgctg	tgccttctag	ttgccagcca	tctgttggtt	gccccctccc	cgtgccttcc	3840
ttgacctggg	aaggtgccac	tcccactgtc	ctttccta	aaaatgagga	aattgcatcg	3900
cattgtctga	gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	3960
gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	gggtaccag	4020
gtgctgaaga	attgaccggg	ttcctcctgg	gccagaaaga	agcaggcaca	tccccctctc	4080
tgtgacacac	cctgtccacg	cccctgggtt	ttagttccag	ccccactcat	aggacactca	4140
tagctcagga	gggctccgcc	ttcaatecca	cccgttaaag	tacttgagac	ggtctctccc	4200
tccctcatca	gccacccaaa	ccaaacctag	cctccaagag	tgggaagaaa	ttaaagcaag	4260
ataggctatt	aagtgcagag	ggagagaaaa	tgctccaac	atgtgaggaa	gtaatgagag	4320
aaatcataga	atthtaaggc	catgatttaa	ggccatcatg	gccttaatct	tccgcttcc	4380
cgctcactga	ctcgctgcgc	tcggctcggtt	ggctgcggcg	agcgggtatca	gctcactcaa	4440
aggcggtaat	acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	4500
aaggccagca	aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcggtt	ttccataggc	4560
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagagggtg	cgaaaccgga	4620
caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	4680
cgacctgccc	gcttaccgga	tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	4740
ctcatagctc	acgctgtagg	tatctcagtt	cgggtgtaggt	cgttcgctcc	aagctgggct	4800
gtgtgcacga	acccccggtt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	4860
agtccaaccc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	4920
gcagagcgag	gtatgtaggc	ggtgctacag	agttcttgaa	gtggtggcct	aactacggct	4980
acactagaag	aacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	5040
gagttggtag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgt	ttttttgttt	5100
gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	5160
cggggtctga	cgctcagtgg	aacgaaaact	cacgttaagg	gatttttggtc	atgagattat	5220
caaaaaggat	cttcacctag	atccttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	5280
gtatatatga	gtaaaacttg	tctgacagtt	accaatgctt	aatcagttag	gcacctatct	5340
cagcgatctg	tctatthcgt	tcattccatag	ttgcctgact	cggggggggg	gggcgctgag	5400
gtctgcctcg	tgaagaagg	gttgcctgact	cataccaggc	ctgaatcgcc	ccatcatcca	5460
gccagaaagt	gaggagagca	cggttgatga	gagctttggt	gtaggtggac	cagttggtga	5520
ttttgaactt	ttgctttgcc	acggaacgg	ctgcgttggtc	gggaagatgc	gtgatctgat	5580
ccttcaactc	agcaaaagt	cgatttatte	aacaaagccg	ccgtcccgtc	aagtcagcgt	5640
aatgctctgc	cagtgttaca	accaattaac	caattctgat	tagaaaaact	catcgagcat	5700
caaatgaaac	tgcaatttat	tcatatcagg	attatcaata	ccatattttt	gaaaaagccg	5760
tttctgtaat	gaaggagaaa	actcaccgag	gcagttccat	aggatggcaa	gatcctggta	5820
tcggctctgc	attccgactc	gtccaacatc	aatacaacct	attaatttcc	cctcgtcaaa	5880
aataaggtta	tcaagtgaga	aatcaccatg	agtgacgact	gaatccggtg	agaatggcaa	5940
aagcttatgc	atttctttcc	agacttggtc	aacaggccag	ccattacgct	cgtcatcaaa	6000
atcactcgca	tcaaccaaac	cgttatccat	tcgtgattgc	gcctgagcga	gacgaaatac	6060
gcgatcgctg	ttaaaaggac	aattacaac	aggaatcgaa	tgcaaccggc	gcaggaacac	6120
tgccagcgca	tcaacaatat	tttcacctga	atcaggatat	tcttctaata	cctggaatgc	6180
tgttttcccc	gggatcgag	tggtagtaa	ccatgcatca	tcaggagtag	ggataaaaatg	6240
cttgatgggtc	ggaagaggca	ttaattccgt	cagccagttt	agtctgacca	tctcatctgt	6300
aacatcattg	gcaacgctac	ctttgccatg	tttcagaaac	aactctggcg	catcgggctt	6360
cccatacaat	cgatagattg	tcgcacctga	ttgcccagca	ttatcgcgag	cccatttata	6420
cccataataa	tcagcatcca	tgttggaatt	taatcgcggc	ctcgagcaag	acgtttcccc	6480

```

ttgaatatgg ctcataacac cccttgatatt actggttatg taagcagaca gttttattgt 6540
tcatgatgat atatTTTTat cttgtgcaat gtaacatcag agatTTTtgag acacaacgtg 6600
gctttccccc cccccccatt attgaagcat ttatcagggg tattgtctca tgagcggata 6660
catatttgaa tgtattttaga aaaataaaca aatagggggt ccgcgcacat ttccccgaaa 6720
agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg 6780
tatcacgagg ccctttcgtc                                     6800

```

&lt;210&gt; 94

&lt;211&gt; 6872

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 94

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgccc aacgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagccctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctccctaacg tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacgc gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgcgt gaaggagacc cagatgaact ggcctaacct gtggaagtgg ggcacccctga 1980
tcctgggcct ggtgatcatc tgcagcgcca gcgacaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggcgcgac gccgacacca ccctgttctg cgccagcgac gccaaaggccc 2100
acgagaccga ggtgcacaac gtgtgggcca cccagcctcg cgtgcccacc gacccaacc 2160
cccaggaggt ggacctggag aacatcaccc agaacttcaa catgtggaag aacaacatgg 2220

```



tggagcagat	gcaggaggac	gtgatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgaact	gcataaacgc	caacctgacc	aacgtgaaca	2340
acatcatcca	cggccccaac	atcatcggca	acatcaccca	cgagggtgcgc	aactgcagct	2400
tcaacatgac	caccgagatc	cgcgacaaga	agcagaaggt	gcacgccctg	ttctacaagc	2460
tggacctggt	gcagatcgag	gacaagaaca	gcagcgagta	ccgcctgatc	aactgcaaca	2520
ccagcgtgat	caagcaggcc	tgccccaaga	tcagcttcga	ccccatcccc	atccactact	2580
gcacccccgc	cggctacgcc	atcctgaagt	gcaacgacaa	gaacttcaac	ggcaccggcc	2640
cctgcaccaa	cgtgagcagc	gtgcagtgc	cccacggcat	caagcccgtg	gtgagcacc	2700
agctgctgct	gaacggcagc	ctggccgagg	aggagatcat	catccgcagc	gagaacctga	2760
ccaacaacgc	caagaccatc	atcgtgcacc	tgaacaagag	cgtggagatc	aactgcaccc	2820
gccccagcaa	caacacccgc	accagcacca	ccatcggccc	cggccagggtg	ttctaccgca	2880
ccggcgacat	caccggcaac	atccgcaagg	cctactgcga	gatcaacggc	accaagtgg	2940
acgaggccct	gaagcagggtg	accaagaagc	tgaaggagca	cttcaacaac	aagaccatca	3000
tcttccagcc	ccccagcggc	ggcgacctgg	agatcaccat	gcaccacttc	aactgccgcg	3060
gcgagttctt	ctactgcaac	accacccgcc	tgttcaacaa	cacctgcata	ggcaacgaga	3120
ccatcggcgg	ctgcaacgac	accatcatcc	tgccctgcaa	gatcaagcag	atcatcaaca	3180
tgtggcaggg	cgccggcctg	gccatgtacg	cccccccat	cagcggccgc	atcaactgcg	3240
tgagcaacat	caccggcatc	ctgctgaccc	gcgacggcgg	cgccaacaac	accagcaacg	3300
agatcttccg	ccccggcggc	ggcaacatca	aggacaactg	gcgcagcgag	ctgtacaagt	3360
acaagggtgt	gcagatcgag	cccctgggca	tcgcccccac	ccgcgccaaag	cttaccgtgc	3420
aggcccgcca	gctggtgagc	ggcatcgtgc	agcagcagag	caacctgctg	cgcgccatcg	3480
aggcccagca	gcacctgctg	cagctgaccg	tgtggggcat	caagcagctg	caggcccgcg	3540
tgctggccgt	ggagcgctac	ctgaaggacc	agaagttcct	cgaggagatc	tggagcaaca	3600
tgacctggat	cgagtgggag	cgcgagatca	gcaactacac	caaccagatc	tacgagatcc	3660
tgaccgagag	ccagaaccag	caggaccgca	acgagaagga	cctgctggag	ctggacaagt	3720
gggcgagcct	gtggagctgg	ttcgacatca	ccaactggct	gtggtacatc	aagatcttca	3780
tcatgatcgt	gggcggcctg	atcggcctgc	gcatactctt	cgcctgctg	agcatctgaa	3840
cacgtgggat	ccagatctgc	tgtgccttct	agttgccagc	catctgttgt	ttgccccctc	3900
cccgtgcctt	ccttgaccct	ggaagggtgc	actcccactg	tccttttcta	ataaaatgag	3960
gaaattgcat	cgcattgtct	gagtaggtgt	cattctatct	tgggggggtg	ggtggggcag	4020
gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	ggtgggctct	4080
atgggtaccc	aggtgctgaa	gaattgaccc	ggttcctcct	gggccagaaa	gaagcaggca	4140
catccccttc	tctgtgacac	accctgtcca	cgcctcctgt	tcttagttcc	agccccactc	4200
ataggacact	catagctcag	gagggtctcg	ccttcaatcc	caccgcctaa	agtacttgga	4260
gcggtctctc	cctccctcat	cagcccacca	aaccaaacct	agcctccaag	agtgggaaga	4320
aattaaagca	agataggcta	ttaagtgcag	aggagagaaa	aatgcctcca	acatgtgagg	4380
aagtaatgag	agaaatcata	gaattttaag	gccatgatct	aaggccatca	tggccttaat	4440
cttccgcttc	ctcgtctact	gactcgtctg	gctcggctcg	tcggctgcgg	cgagcgggat	4500
cagctcactc	aaaggcggta	atacggttat	ccacagaatc	aggggataac	gcaggaaaga	4560
acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	4620
ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	4680
ggcgaaaccc	gacaggacta	taaagatacc	aggcgtttcc	ccctggaagc	tccctcgtgc	4740
gctctcctgt	tccgaccctg	ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	4800
gcgtggcgct	ttctcatagc	tcacgctgta	ggtatctcag	ttcgggtgtag	gtcgttcgct	4860
ccaagctggg	ctgtgtgcac	gaaccccccg	ttcagccccg	ccgctgcgcc	ttatccggta	4920
actatcgtct	tgagtccaac	ccggtaagac	acgacttatc	gccactggca	gcagccactg	4980
gtaacaggat	tagcagagcg	aggtatgtag	gcggtgctac	agagtctctg	aagtgggtggc	5040
ctaactacgg	ctacactaga	agaacagtat	ttggtatctg	cgtctgctg	aagccagtta	5100
ccttcggaaa	aagagtgtgt	agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggtg	5160
gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	5220
tgatcttttc	tacgggggtct	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	5280
tcatgagatt	atcaaaaagg	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	5340
aatcaatcta	aagtatatat	gagtaaacct	ggtctgacag	ttaccaatgc	ttaatcagtg	5400
aggcacctat	ctcagcgatc	tgtctatttc	gttcatccat	agttgcctga	ctcggggggg	5460
gggggcgctg	aggtctgcct	cgtgaagaag	gtgttgctga	ctcataccag	gcctgaatcg	5520

```

ccccatcatc cagccagaaa gtgaggggagc cacgggttgat gagagctttg ttgtaggtgg 5580
accagttggt gatcttgaac ttttgctttg ccacggaacg gtctgcgttg tcgggaagat 5640
gcgtagatctg atccttcaac tcagcaaaaag ttcgatttat tcaacaaagc cgccgtcccg 5700
tcaagtcagc gtaatgctct gccagtggtta caaccaatta accaattctg attagaaaaa 5760
ctcatcgagc atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt 5820
ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc 5880
aagatcctgg tatcgggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt 5940
cccctcgta aaaataaggt tatcaagtga gaaatcacca tgagtgaaga ctgaatccgg 6000
tgagaatggc aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacg 6060
ctcgtcatca aaatcactcg catcaacca accgttattc attcgtgatt gcgcctgagc 6120
gagacgaaat acgcatcg cgtttaaagg acaattacaa acaggaatcg aatgcaaccg 6180
gcgtaggaac actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa 6240
tacctggaat gctgttttcc cggggatcg agtggtgagt aaccatgcat catcaggagt 6300
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 6360
catctcatct gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg 6420
cgcacgggc ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg 6480
agccatttta taccatata aatcagcatc catgttgga tttaatcgcg gcctcgagca 6540
agacgtttcc cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga 6600
cagttttatt gttcatgatg atatatTTTT atcttgtgca atgtaacatc agagattttg 6660
agacacaacg tggctttccc cccccccca ttattgaagc atttatcagg gttattgtct 6720
catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 6780
atctcccgga aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta 6840
taaaaatagg cgtatcacga ggccctttcg tc 6872

```

&lt;210&gt; 95

&lt;211&gt; 6757

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 95

```

tcgcgcgttt cggtagatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc accgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgectg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca ccccttgggc 1080
tcttatgcat gctatactgt ttttggttgg ggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200

```



tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttacaattt	cacatatata	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcggttac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgtcttaga	gatatcgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgttaact	ggccgcagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggg	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gtccttcaa	caccaccacc	gaaatccgtg	2400
acaaaaaaca	gcagggttac	gctctgttct	accgtccgga	catcgttctg	ctgaaagaaa	2460
accgtaacaa	ctccaacaac	tccgaatata	tcctgatcaa	ctgcaacgct	tccaccatca	2520
cccaggcttg	cccgaaggtt	aacttcgacc	cgatcccgat	ccactactgc	gctccggctg	2580
gttacgctat	cctgaaatgc	aacaacaaaa	ccttctccgg	taaagggtccg	tgcaacaacg	2640
tttccaccgt	tcagtgcacc	catggtatca	aaccggttgt	ttccaccag	ctgctgctga	2700
acggttccct	ggctgaaaaa	gaaatcatca	tcggttccga	aaacctgacc	gacaacgtta	2760
aaaccatcat	cgttcacctg	aacaaatccg	ttgaaatcgt	ttgcacccgt	ccgaacaaca	2820
acacccgtaa	atccatgcgt	atcgggtccg	gtcagacctt	ctacgctacc	ggtgacatca	2880
tcggtgacat	ccgtcaggct	tactgcaaca	tctccgggtc	caaagtgaac	gaaaccctga	2940
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	3000
cgtcctccgg	tgggtgacctg	gaaatcacca	cccactcctt	caactgccgt	ggtgaattct	3060
tctactgcaa	caccaccggt	ctgttcaaca	acaacgctac	cgaagacgaa	accatcaccc	3120
tgccgtgccg	tatcaaacag	atcatcaaca	tgtggcaggg	tgttgggtcgt	gctatgtacg	3180
ctccgccgat	cgttggtaac	atcacctgca	aatccaacat	caccggtctg	ctgctgggtc	3240
gtgacgggtg	tgaagacaac	aaaaccgaag	aaatcttccg	tccgggtggg	ggtaacatga	3300
aagacaactg	gcgttccgaa	ctgtacaaat	acaaagtatt	cgaactgaaa	ccgctgggta	3360
tcgctccgac	cgggtgctaag	cttaccgttc	aggctcgtca	gctgctgtcc	tccatcggtc	3420
agcagcagtc	caacctgctg	cgtgctatcg	aaagctcagca	gcacatgctg	cagctgaccg	3480
tttgggggtat	caaacagctg	cagaccgctg	ttctggctat	cgaacgttac	ctgaaagacc	3540
agcagctcga	gatctggaac	aacatgacct	ggatggaatg	ggaccgtgaa	atctccaact	3600
acaccgacac	catctaccgt	ctgctggaag	actcccagac	ccagcaggaa	aaaaacgaaa	3660
aagacctgct	ggctctggac	tcttgaaaaa	acctgtgggtc	ctggttcgac	atctccaact	3720
ggctgtgggtg	aggatccaga	tctgctgtgc	cttctagttg	ccagccatct	gttggtttgcc	3780
cctcccccggt	gccttccttg	accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	3840
atgaggaaat	tgcategcat	tgtctgagta	gggtgtcatc	tattctgggg	ggtgggggtg	3900
ggcaggacag	caagggggag	gattgggaag	acaatagcag	gcatgctggg	gatgcgggtg	3960
gctctatggg	taccaggtg	ctgaagaatt	gacccggttc	ctcctggggc	agaaagaagc	4020
aggcacatcc	ccttctctgt	gacacacctt	gtccacgccc	ctggttctta	gttccagccc	4080
cactcatagg	acactcatag	ctcaggaggg	ctccgccttc	aatcccaccc	gctaaagtac	4140
ttggagcggt	ctctccctcc	ctcatcagcc	caccaaacca	aacctagcct	ccaagagtgg	4200
gaagaaatta	aagcaagata	ggctattaag	tgcagaggga	gagaaaatgc	ctccaacatg	4260
tgaggaagta	atgagagaaa	tcatagaatt	ttaaggccat	gatttaaggc	catcatggcc	4320
ttaatcttcc	gcttcctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	tgccggcgagc	4380
ggtatcagct	cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	4440
aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccagggaac	cgtaaaaaagg	ccgcgttgct	4500

```

ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4560
gaggtggcga aaccgcagag gactataaag ataccaggcg ttccccctg gaagctccct 4620
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4680
gggaagcgtg gcgctttctc atagctcacg ctgtagggtat ctcagttcgg tgtaggtcgt 4740
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttate 4800
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 4860
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4920
gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 4980
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 5040
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 5100
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5160
tttggtcatg agattatcaa aaaggatctt cacctagatc ctttttaaatt aaaaatgaag 5220
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5280
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5340
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5400
aatcgcccca tcatccagcc agaaagttag ggagccacgg ttgatgagag ctttgttgta 5460
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacgggtctg cgttgtcggg 5520
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5580
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5640
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5700
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5760
atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5820
aatttccctt cgtcaaaaat aagggttatca agtgagaaat caccatgagt gacgactgaa 5880
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5940
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 6000
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 6060
aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc aggatattct 6120
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcacatca 6180
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6240
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6300
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6360
tcgcgagccc atttataccc atataaatca gcattccatgt tggaaatttaa tcgcggcctc 6420
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6480
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6540
ttttgagaca caacgtggct tccccccccc cccattatt gaagcattta tcagggttat 6600
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6660
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6720
acctataaaa ataggcgat cagaggccc tttcgtc

```

&lt;210&gt; 96

&lt;211&gt; 6829

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 96

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300

```

```

tccaacatta cgcgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcggt acataactta cggtaaatgg 420
cccgccctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatgggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tctttcccat gggcttttcc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcacc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca cctgttcttg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gtccttcaa caccaccacc gaaatccgtg 2400
acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460
accgtaacaa ctccaacaac tccgaataca tctgatcaa ctgcaacgct tccaccatca 2520
cccaggcttg ccgaaagtt aacttcgacc cgatcccgat ccaactactg gctccggctg 2580
gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaagggtccg tgcaacaacg 2640
tttccaccgt tcagtgcacc catggtatca aaccggttgt ttccaccag ctgctgctga 2700
acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760
aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcacctg ccgaacaaca 2820
acaccgtaa atccatgcgt atcgggtccg gtcagacct ctacgctacc ggtgacatca 2880
tcggtgacat ccgtcaggct tactgcaaca tctccggttc caaatggaac gaaacctga 2940
aacgtgttaa agaaaaactg caggaaaact acaacaacaa caaaaccatc aaattcgctc 3000
cgtcctccgg tggtagacct gaaatcacca cccactcctt caactgccgt ggtgaattct 3060
tctactgcaa caccaccgct ctgttcaaca acaacgctac cgaagacgaa accatcacc 3120
tgccgtgccg tatcaaacag atcatcaaca tgtggcaggg tgttggtcgt gctatgtacg 3180
ctccgccgat cgtggtaac atcacctgca aatccaacat caccggtctg ctgctggttc 3240
gtgacgggtg tgaagacaac aaaaccgaag aaatcttccg tccgggtggg ggtaacatga 3300
aagacaactg gcgttccgaa ctgtacaaat acaaagttat cgaactgaaa ccgctgggta 3360
tcgtcccgac cgtgtctaag ctacccgttc aggtcgtca gctgctgtcc tccatcgctc 3420
agcagcagtc caacctgctg cgtgctatcg aagctcagca gcacatgctg cagctgaccg 3480
tttgggggtat caaacagctg cagaccctgt ttctggctat cgaacgttac ctgaaagacc 3540
agcagctcga gatctggaac aacatgacct ggatggaatg ggaccgtgaa atctccaact 3600

```



acaccgacac	catctaccgt	ctgctggaag	actcccagac	ccagcaggaa	aaaaacgaaa	3660
aagacctgct	ggctctggac	tcctggaaaa	acctgtggtc	ctggttcgac	atctccaact	3720
ggctgtggta	catcaaaatc	ttcatcatga	tcgttggtgg	tctgatcggg	ctgcgtatca	3780
tcttcgctgt	tctgtccatc	tgaggatcca	gatctgctgt	gccttctagt	tgccagccat	3840
ctgttggttg	cccctcccc	gtgccttcc	tgaccctgga	aggtgccact	cccactgtcc	3900
tttccctaata	aaatgaggaa	attgcatcgc	attgtctgag	taggtgtcat	tctattctgg	3960
ggggtggggg	ggggcaggac	agcaaggggg	aggattggga	agacaatagc	aggcatgctg	4020
gggatgcggg	gggctctatg	ggtacccagg	tgctgaagaa	ttgaccggg	tcctcctggg	4080
ccagaaagaa	gcaggcacat	ccccttctct	gtgacacacc	ctgtccacgc	ccctgggttct	4140
tagttccagc	cccactcata	ggacactcat	agctcaggag	ggctccgcct	tcaatcccac	4200
ccgctaaagt	acttgagcgc	gtctctccct	ccctcatcag	cccaccaaac	caaacctagc	4260
ctccaagagt	gggaagaaat	taaagcaaga	taggctatta	agtgcagagg	gagagaaaat	4320
gcctccaaca	tgtgaggaag	taatgagaga	aatcatagaa	ttttaaggcc	atgatttaag	4380
gccatcatgg	ccttaatctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggtcgttcg	4440
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggttaata	cggttatcca	cagaatcagg	4500
ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	4560
ggccgcgttg	ctggcgtttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	4620
acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	4680
tggaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	4740
ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	cgctgtagg	atctcagttc	4800
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggttc	agcccgaccg	4860
ctgcgcctta	tccggttaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	4920
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	4980
gttcttgaag	tggtagccta	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	5040
tctgctgaag	ccagttacct	tccgaaaaag	agttggtagc	tcttgatccg	gcaaaacaaac	5100
caccgctggg	agcgggtggt	tttttggttg	caagcagcag	attacgcgca	gaaaaaaagg	5160
atctcaagaa	gatcctttga	tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	5220
acgttaaggg	attttggtca	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	5280
ttaaaaatga	agttttaaat	caatctaaag	tatatatgag	taaacttggg	ctgacagtta	5340
ccaatgctta	atcagtgagg	cacctatctc	agcgatctgt	ctatttcgtt	catccatagt	5400
tgctgactc	gggggggggg	ggcgctgagg	tctgcctcgt	gaagaagggtg	ttgctgactc	5460
ataccaggcc	tgaatcgccc	catcatccag	ccagaaagtg	aggagaccac	ggttgatgag	5520
agctttggtg	taggtggacc	agttgggtgat	tttgaacttt	tgctttgcca	cggaacggtc	5580
tgcgttgtcg	ggaagatgcg	tgatctgatc	cttcaactca	gcaaaaagttc	gattttattca	5640
acaaagccgc	cgtcccgtca	agtcagcgta	atgctctgcc	agtgttacia	ccaattaacc	5700
aattctgatt	agaaaaactc	atcgagcatc	aaatgaaact	gcaatttatt	catatcagga	5760
ttatcaatac	catatttttg	aaaaagccgt	ttctgtaatg	aaggagaaaa	ctcaccgagg	5820
cagttccata	ggatggcaag	atcctgggtat	cggctctgca	ttccgactcg	tccaacatca	5880
atacaacctc	ttaatttccc	ctcgtcaaaa	ataagggtat	caagtggaga	atcaccatga	5940
gtgacgactg	aatccgggtga	gaatggcaaa	agcttatgca	tttctttcca	gacttggtca	6000
acaggccagc	cattacgctc	gtcatcaaaa	tcactcgcat	caaccaaacc	gttattcatt	6060
cgtgattgcg	cctgagcgag	acgaaatacg	cgatcgctgt	taaaaggaca	attacaaaca	6120
ggaatcgaat	gcaaccggcg	caggaacact	gccagcgcat	caacaatatt	ttcacctgaa	6180
tcaggatatt	cttctaatac	ctggaatgct	gttttcccgg	ggatcgcagt	ggtgagtaac	6240
catgcatcat	caggagtacg	gataaaatgc	ttgatggctg	gaagaggcat	aaattccgctc	6300
agccagttta	gtctgaccat	ctcatctgta	acatcattgg	caacgctacc	tttgccatgt	6360
ttcagaaaca	actctggcgc	atcgggcttc	ccatacaatc	gatagattgt	cgcacctgat	6420
tgcccagacat	tatcgcgagc	ccatttatac	ccatataaat	cagcatccat	gttgggaattt	6480
aatcgcgggc	tcgagcaaga	cgtttcccgt	tgaatatggc	tcataacacc	ccttggtatta	6540
ctgtttatgt	aagcagacag	ttttattggt	catgatgata	tatttttata	ttgtgcaatg	6600
taacatcaga	gattttgaga	cacaacgtgg	ctttcccccc	ccccccatta	ttgaagcatt	6660
tatcagggtt	attgtctcat	gagcggatag	atatttgaat	gtatttagaa	aaataaacia	6720
ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	acgtctaaga	aaccattatt	6780
atcatgacat	taacctataa	aaataggcgt	atcacgaggc	cctttcgtc		6829

&lt;210&gt; 97

&lt;211&gt; 6752

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 97

```
tcgcgcggttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc ggcatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcgg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcggttac gtgttccgga 1440
catgggctct tctccggtag cggcggtgac tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggcttttcc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagt gtggatctgg ggcaccttg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgaagaac tgcagcttca 2340
acatgaccac cgaggtgcgc gacaagcgcg agaaggtgca cgccctgttc taccgcctgg 2400
acatcgtgcc cctgaagaac gagagcagca acaccagcg cgactaccgc ctgatcaact 2460
gcaacaccag cgcacatcacc caggcctgcc ccaaggtgag cttcgacccc atccccatcc 2520
actactgcgc ccccgccggc tacgccatcc tgaagtgcaa caacaagacc ttcaacggca 2580
ccggcccttg caacaacgtg agcaccatcc agtgcacca cggcaccaag cccgtggtga 2640
```



gcacccagct	gctgctgaac	ggcagcctgg	ccgaggagga	gatcatcatc	cgcagcaaga	2700
acctgaccga	caacgtgaag	accatcatcg	tgcacctgaa	cgagagcgtg	gagatcaact	2760
gcacccgccc	caacaacaac	acccgcaaga	gcateccgat	cggccccggc	caggccttct	2820
acgccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacatc	agccgcaccg	2880
cctggaacaa	gaccttgacg	gaggtgggca	agaagctggc	cgagcacttc	cccaacaagg	2940
ccatcaagtt	cgccaagcac	agcggcgggc	acctggagat	caccacccac	agcttcaact	3000
gccgcggcga	gttcttctac	tgcaacacca	gcagcctggt	caacagcacc	tacaccccca	3060
acagcaccga	gaacatcacc	ggcaccgaga	acagcatcat	caccatcccc	tgccgcacat	3120
agcagatcat	caacatgtgg	caggggcgtg	gccgcggccat	gtacgcccc	cccatcgagg	3180
gcatectgac	ctgccgcagc	aacatcaccg	gcctgctgct	gacccgcgac	ggcggcaccg	3240
gcatgcacga	caccgagatc	ttccgccccg	agggcgggca	catgcgcgac	aactggcgca	3300
gcgagctgta	caagtacaag	gtggtggaga	tcaagccctt	gggcatcgcc	cccaccaagg	3360
ccaagcttac	cgtgcaggtg	cgccagctgc	tgagcggcat	cgtgcagcag	cagagcaacc	3420
tgctgcgcgc	catcgaggcc	cagcagcaca	tgctgcagct	gaccgtgtgg	ggcatcaagc	3480
agctgcagac	ccgcgtgctg	gccatcgagc	gctacctgcy	cgaccagcag	ctcgaggaca	3540
tctggaacaa	catgacctgg	atgcagtggg	accgcgagat	cagcaactac	accaacacca	3600
tctaccgcct	gctggaggac	agccagaacc	agcaggagaa	gaacgagcag	gacctgctgg	3660
ccctggacaa	gtggcagaac	ctgtggacct	ggttcggcat	caccaactgg	ctgtggtgaa	3720
cacgtgggat	ccagatctgc	tgtgccttct	agttgccagc	catctgttgt	ttgcccctcc	3780
cccgtgcctt	ccttgaccct	ggaaggtgcc	actcccactg	tcctttccta	ataaaatgag	3840
gaaattgcat	cgcatttgtct	gagtaggtgt	cattctatct	tgggggggtg	ggtggggcag	3900
gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctgggggatg	ggtgggctct	3960
atgggtaccc	aggtgctgaa	gaattgacct	ggttcctcct	gggccagaaa	gaagcaggca	4020
catccccctc	tctgtgacac	accctgtcca	cgccccgtgt	tcttagttcc	agccccactc	4080
ataggacact	catagctcag	gagggctccg	ccttcaatcc	cacccgctaa	agtacttgga	4140
gcggtctctc	cctccctcat	cagcccacca	aaccaaacct	agcctccaag	agtgggaaga	4200
aattaaagca	agataggcta	ttaagtgcag	aggagagaaa	aatgcctcca	acatgtgagg	4260
aagtaatgag	agaaatcata	gaattttaag	gccatgatbt	aaggccatca	tggccttaat	4320
cttccgcttc	ctcgctcact	gactcgctgc	gctcggctcg	tccgctgcgg	cgagcgggat	4380
cagctcactc	aaaggcggtg	atacggttat	ccacagaatc	aggggataac	gcaggaaaaga	4440
acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcy	ttgctggcgt	4500
ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	agtcagagggt	4560
ggcgaaaccc	gacaggacta	taaagatacc	aggcgtttcc	ccctggaage	tccctcgtgc	4620
gctctcctgt	tccgaccctg	ccgcttaccc	gatacctgtc	cgcctttctc	ccttcgggaa	4680
gcgtggcgct	ttctcatagc	tcacgctgta	ggtatctcag	ttcgggtgtg	gtcgttcgct	4740
ccaagctggg	ctgtgtgcac	gaaccccccg	ttcagcccca	ccgctgcgcc	ttatccggta	4800
actatcgctc	tgagtccaac	ccggtaagac	acgacttctc	gccactggca	gcagccactg	4860
gtaacaggat	tagcagagcg	aggtatgtag	gcggtgctac	agagtctctg	aagtgggtggc	4920
ctaactacgg	ctacactaga	agaacagtat	ttggtatctg	cgctctgctg	aagccagtta	4980
ccttcggaaa	aagagttggg	agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggtg	5040
gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	5100
tgatcttttc	tacgggggtc	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	5160
tcatgagatt	atcaaaaagg	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	5220
aatcaatcta	aagtatatat	gagtaaactt	ggtctgacag	ttaccaatgc	ttaatcagtg	5280
aggcacctat	ctcagcgatc	tgtctatttc	gttcatccat	agttgcctga	ctcggggggg	5340
gggggcgctg	aggtctgcct	cgtgaagaag	gtgttgctga	ctcataccag	gcctgaatcg	5400
ccccatcatc	cagccagaaa	gtgaggagac	cacggttgat	gagagctttg	ttgtagggtg	5460
accagttggg	gattttgaac	ttttgctttg	ccacggaacg	gtctgcgttg	tcgggaagat	5520
gcgtgatctg	atccttcaac	tcagcaaaaag	ttcgatttat	tcaacaaage	cgcctgcccc	5580
tcaagtcagc	gtaatgctct	gccagtgtta	caaccaatta	accaattctg	attagaaaaa	5640
ctcatcgagc	atcaaatgaa	actgcaattt	attcatatca	ggattatcaa	taccatattt	5700
ttgaaaaagc	cgtttctgta	atgaaggaga	aaactcaccg	aggcagttcc	ataggatggc	5760
aagatccctg	tatcggctctg	cgattccgac	tcgtccaaca	tcaatacaac	ctattaattt	5820
cccctcgtca	aaaataagggt	tatcaagtga	gaaatcacca	tgagtgcaga	ctgaatccgg	5880
tgagaatggc	aaaagcttat	gcatttcttt	ccagacttgt	tcaacaggcc	agccattacg	5940

```

ctcgtcatca aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc 6000
gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg 6060
gcgcaggaac actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa 6120
tacctggaat gctgttttcc cggggatcgc agtgggtgagt aaccatgcat catcaggagt 6180
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 6240
catctcatct gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg 6300
cgcatcgggc ttcccatata atcgatagat tgtcgcacct gattgcccga cattatcgcg 6360
agcccattta taccatata aatcagcatc catgtttgga tttaatcgcg gcctcgagca 6420
agacgtttcc cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga 6480
cagttttatt gttcatgatg atatatTTTT atcttgtgca atgtaacatc agagattttg 6540
agacacaacg tggctttccc cccccccca ttattgaagc atttatcagg gttattgtct 6600
catgagcgga tacatatTTg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 6660
atttccccga aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaacct 6720
taaaaatagg cgtatcacga ggccctttcg tc 6752

```

&lt;210&gt; 98

&lt;211&gt; 6623

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 98

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgtgttt ttgacctcca 960
tagaagacac cgggaccgat ccagccctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggtttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

```

gttgccggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcattctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgctgta 2280
agctgacccc cctgtgctg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340
gcgccatcac ccaggcctgc cccaagggtga gcttcgaccc catccccatc cactactgcg 2400
ccccgcgcgg ctacgccatc ctgaagtga acaacaagac cttcaacggc accggccctc 2460
gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccggtggtg agcaccagc 2520
tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcaccgccc 2640
ccaacaacaa caccgcaag agcatccgca tcggccccgg ccaggccttc tacgccaccg 2700
gcgagatcat cggcgacatc cgccaggccc actgcaacat cagccgcacc gcctggaaca 2760
agaccctgca ggaggtgggc aagaagctgg ccgagcactt cccaacaag gccatcaagt 2820
tcgccaagca cagcggcggc gacctggaga tcaccacca cagcttcaac tgccgcggcg 2880
agttcttcta ctgcaacacc agcagcctgt tcaacagcac ctacaccccc aacagcaccg 2940
agaacatcac cggcaccgag aacagcatca tcaccatccc ctgccgcac aagcagatca 3000
tcaacatgtg gcagggcggt ggccgcgcca tgtacgcccc ccccatcgag ggcattctga 3060
cctgccgcag caacatcac ggctgctgc tgaccgcga cggcggcacc ggcattgcag 3120
acaccgagat ctccgcccc gagggcgcg acatgcgga caactggcg agcgagctgt 3180
acaagtacaa ggtggtggag atcaagcccc tgggcatcgc cccaccaag gccaaagctta 3240
ccgtgcaggt gcgccagctg ctgagcggca tcgtgcagca gcagagcaac ctgctgcgcg 3300
ccatcgaggc ccagcagcac atgctgcagc tgaccgtgtg gggcatcaag cagctgcaga 3360
cccgctgct ggccatcgag cgtacctgc gcgaccagca gctcgaggac atctggaaca 3420
acatgacctg gatgcagtg gaccgcgaga tcagcaacta caccaacacc atctaccgcc 3480
tgctggagga cagccagaac cagcaggaga agaacgagca ggacctgctg gccctggaca 3540
agtggcagaa cctgtggacc tggttcggca tcaccaactg gctgtggtga acacgtggga 3600
tccagatctg ctgtgccttc tagttgccag ccatctgttg tttgccccct ccccgctgcct 3660
tccttgaccc tggaaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca 3720
tcgcattgtc tgagtaggtg tcattctatt ctgggggggt ggggtggggca ggacagcaag 3780
ggggaggatt ggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgggtacc 3840
caggtgctga agaattgacc cggttcctcc tggggcagaa agaagcaggc acatccccct 3900
ctctgtgaca caccctgtcc acgccccctg ttcttagttc cagccccact cataggacac 3960
tcatagctca ggagggctcc gccttcaatc ccaccgccta aagtacttgg agcggctctc 4020
ccctccctca tcagcccacc aaaccaaac tagcctccaa gagtgggaag aaattaaagc 4080
aagataggct attaatgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4140
gagaaatcat agaattttta ggccatgatt taaggccatc atggccttaa tcttcgctt 4200
cctcgctcac tgactcgctg cgtcggctcg ttccggctgc gcgagcggta tcagctcact 4260
caaaggcgg aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag 4320
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 4380
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 4440
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg 4500
ttccgacct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcg 4560
tttctcatag ctacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg 4620
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccgg aactatcgtc 4680
ttgagtccaa cccggttaaga cagcacttat cgccactggc agcagccact ggtaacagga 4740
ttagcagagc gaggtatgta ggcggtgcta cagagtctct gaagtgggtg cctaactacg 4800
gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 4860
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg 4920
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 4980
ctacgggggtc tgacgctcag tggaaacgaaa actcacgtta agggattttg gtcattgagat 5040

```



```

tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 5100
aaagtatatata tgagtaaact tgggtctgaca gttaccaatg cttaatcagt gaggcaccta 5160
tctcagcgat ctgtctattt cgttcattcca tagttgcctg actcgggggg ggggggcgct 5220
gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc gcccattcat 5280
ccagccagaa agtgaggagg ccacgggtga tgagagcttt gttgtagggtg gaccagttgg 5340
tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct 5400
gatccttcaa ctcagcaaaa gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag 5460
cgtaatgctc tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag 5520
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag 5580
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5640
gtatcgggtct gcgattccga ctcggtccaac atcaatacaa cctattaatt tcccctcgtc 5700
aaaaataagg ttatcaagtg agaaatcacc atgagtgcag actgaatccg gtgagaatgg 5760
caaaagctta tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc 5820
aaaatcactc gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa 5880
tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgaggaa 5940
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 6000
tgctgttttc ccggggatcg cagtgggtgag taaccatgca tcatcaggag tacggataaa 6060
atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc 6120
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg 6180
cttcccatac aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt 6240
atacccatat aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgttct 6300
ccgttgataa tggtcataa cacccttgtt attactgttt atgtaagcag acagttttat 6360
tgttcatgat gatataattt tatcttggtgc aatgtaacat cagagatttt gagacacaac 6420
gtggctttcc ccccccccc attattgaag catttatcag gggtattgtc tcatgagcgg 6480
atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 6540
aaaagtgcc cctgacgtct aagaaacatc tattatcatg acattaacct ataaaaatag 6600
gcgtatcacg aggccctttc gtc 6623

```

&lt;210&gt; 99

&lt;211&gt; 6542

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 99

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctggt ttgacctcca 960

```

tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020  
tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctataactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320  
ggatgggggc ccatltatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380  
cgagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500  
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgccgtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgcgtccgcg 1800  
cgccgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctttttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga gatatcgcca 1920  
ccatgcgcgt ggagggcac cagcgcact ggaagcagtg gtggatctgg ggcacccctg 1980  
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040  
gcgtgcccgt gtggaaggag gccaagacca cctgttctg cgcagcgac gccaaggcct 2100  
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160  
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220  
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag cctgctga 2280  
agctgacccc cctgtgctg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340  
gcgccatcac ccaggcctgc cccaaggta gcttcgaccc catccccatc cactactgcg 2400  
cccccgccgg ctacgccatc ctgaagtga acaacaagac cttcaacggc accggccctc 2460  
gcaacaacgt gagcaccatc cagtgcacc acggcaccaa gcccggtggtg agcaccagc 2520  
tgcgtctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580  
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcaccgcc 2640  
gatcgcccca ctgcaacatc agccgcaccg cctggaacaa gacctgcag gaggtgggca 2700  
agaagctggc cgagcacttc cccaacaagg ccatcaagtt cgccaagcac agcggcggcg 2760  
acctggagat caccacccac agcttcaact gccgcggcga gttcttctac tgcaacacca 2820  
gcagcctgtt caacagcacc tacaccccca acagcaccga gaacatcacc ggcaccgaga 2880  
acagcatcat caccatcccc tgcgcacatc agcagatcat caacatgtgg cagggcggtg 2940  
gccgcgccat gtacgcccc cccatcgagg gcacctgac ctgcccagc aacatcaccg 3000  
gcctgctgct gacccgcgac ggcggcaccg gcacgcacga caccgagatc ttccgccccg 3060  
agggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtggaga 3120  
tcaagcccct gggcatcgcc cccaccaagg ccaagcttac cgtgcaggtg cgccagctgc 3180  
tgagcggcat cgtgcagcag cagagcaacc tgcgtgcgcg catcgaggcc cagcagcaca 3240  
tgcgtcagct gaccgtgtgg ggcacaaagc agctgcagac ccgctgctg gccatcgagc 3300  
gctacctgcg cgaccagcag ctgcaggaca tctggaacaa catgacctgg atgcagtggg 3360  
accgcgagat cagcaactac accaacacca tctaccgct gctggaggac agccagaacc 3420  
agcaggagaa gaacgagcag gacctgctgg cctgggacaa gtggcagaac ctgtggacct 3480  
ggttcggcat caccaactgg ctgtggtgaa cacgtgggat ccagatctgc tgtgccttct 3540  
agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaagggtgc 3600  
actcccactg tcccttctta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 3660  
cattctattc tgggggggtg ggtggggcag gacagcaagg gggaggattg ggaagacaat 3720  
agcaggcatg ctggggatgc ggtgggctct atgggtaccc aggtgctgaa gaattgacct 3780  
ggttctctct gggccagaaa gaagcaggca catccccctc tctgtgacac accctgtcca 3840  
cgccccgtg tcttagttcc agccccactc ataggacact catagctcag gagggtccg 3900  
ccttcaatcc caccgcctaa agtacttga gcggtctctc cctccctcat cagccacca 3960  
aaccaaacct agcctccaag agtgggaaga aattaaagca agataggcta ttaagtgcag 4020  
agggagagaa aatgcctcca acatgtgagg aagtaatgag agaaatcata gaattttaag 4080  
gccatgattt aaggccatca tggccttaat ctccgcttc ctcgctcact gactcgtgc 4140  
gctcggctgt tcggctgcgg cgagcggat cagctcactc aaaggcggt atacggttat 4200  
ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca 4260



```

ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc 4320
atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc 4380
aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg 4440
gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta 4500
ggatctctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg 4560
ttcagcccca ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggttaagac 4620
acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag 4680
gcggtgctac agagtctctg aagtgggtgg ctaactacgg ctacactaga agaacagtat 4740
ttggtatctg cgtctcgtg aagccagtta ccttcggaaa aagagttggg agctcttgat 4800
ccggcaaaaa aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc 4860
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc taccgggtct gacgctcagt 4920
ggaacgaaaa ctacgcttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct 4980
agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaacctt 5040
ggctctgacag ttaccaatgc ttaatcagtg aggcacctat ctacgcatc tgtctatttc 5100
gttcatccat agttgcctga ctccgggggg gggggcgctg aggtctgcct cgtgaagaag 5160
gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc 5220
cacggttgat gagagctttg ttgtaggtgg accagttggg gatthttgaa ttttgctttg 5280
ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaaag 5340
ttcgatttat tcaacaaagc cgcctgcccg tcaagtcagc gtaatgctct gccagtgtta 5400
caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt 5460
attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga 5520
aaactcaccg aggcagttcc ataggatggc aagatcctgg taccggtctg cgattccgac 5580
tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt tatcaagtga 5640
gaaatcacca tgagtgcga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt 5700
ccagacttgt tcaacaggcc agccattacg ctctcatca aaatcactcg catcaaccaa 5760
accgttattc attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg 5820
acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg catcaacaat 5880
atthttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc 5940
agtggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg 6000
cataaattcc gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct 6060
acctttgcca tgtttcagaa acaactctgg cgcctcgggc tccccatata atcgatagat 6120
tgtcgcacct gattgcccga cattatcgcg agcccattta taccatata aatcagcatc 6180
catgttgga tttaatcgcg gcctcgagca agacgtttcc cgttgaatat ggctcataac 6240
acccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg atataatttt 6300
atcttggtgca atgtaacatc agagattttg agacacaacg tggctttccc cccccccca 6360
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 6420
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta 6480
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg 6540
tc

```

&lt;210&gt; 100

&lt;211&gt; 6473

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 100

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240

```

ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacattttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaattg	420
cccgctggc	tgaccgcca	acgacccccg	cccattgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgccct	ggcattatgc	ccagtacatg	accttatggg	actttccctac	660
ttggcagtac	atctacgtat	tagtcacgcg	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgtga	780
cgtcaatggg	agtttggttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtatacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccccttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttataaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcgaact	ggaagcagtg	gtggatctgg	ggcactcctgg	1980
gcttctggat	ggtgatgatac	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaagacca	ccctgttctg	cgcagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	atcgatcaac	tgcaacacca	2340
gcgccatcac	ccaggcctgc	cccaagggtga	gcttcgaccc	catccccatc	cactactgcg	2400
cccccgccgg	ctacgccatc	ctgaagtgc	acaacaagac	cttcaacggc	accggccccct	2460
gcaacaacgt	gagcaccatc	cagtgcaccc	acggcaccaa	gcccgtgggtg	agcaccacgc	2520
tgctgctgaa	cggcagcctg	gccgaggagg	agatcatcat	ccgcagcaag	aacctgaccg	2580
acaacgtgaa	gaccatcatc	gtgcacctga	acgagagcgt	ggagatcaac	tgcacccgcc	2640
gacgggcccc	ctgcaacatc	agccgcaccg	cctggaacaa	gaccctgcag	gagggtgggca	2700
agaagctggc	cgagcacttc	cccaacaagg	ccatcaagtt	cgccaagcac	agcggcggcg	2760
acctggagat	caccaccac	agcttcaact	gccgcggcga	gttctttctac	tgcaacaccc	2820
gacgatccc	ctgccgcatac	aagcagatca	tcaacatgtg	gcagggcgtg	ggccgcgcca	2880
tgtacgcccc	ccccatcgag	ggcatcctga	cctgccgcag	caacatcacc	ggcctgctgc	2940
tgacccgcga	cggcggcacc	ggcatgcacg	acaccgagat	cttccgcccc	gagggcggcg	3000
acatgcgcga	caactggcgc	agcgagctgt	acaagtacaa	ggtgggtggag	atcaagcccc	3060
tgggcategc	ccccaccaag	gccaagctta	ccgtgcaggt	gcgccagctg	ctgagcggca	3120
tcgtgcagca	gcagagcaac	ctgctgcgcg	ccatcgaggc	ccagcagcac	atgctgcagc	3180
tgaccgtgtg	gggcatcaag	cagctgcaga	cccgcgtgct	ggccatcgag	cgctacctgc	3240
gcgaccagca	gctcgaggac	atctggaaca	acatgacctg	gatgcagtgg	gaccgcgaga	3300
tcagcaacta	caccaacacc	atctaccgcc	tgctggagga	cagccagaac	cagcaggaga	3360
agaacgagca	ggacctgctg	gccctggaca	agtggcagaa	cctgtggacc	tggttcggca	3420
tcaccaactg	gctgtggtga	acacgtggga	tccagatctg	ctgtgccttc	tagttgccag	3480
ccatctgttg	tttgccccctc	ccccgtgcct	tccttgaccc	tgggaaggtgc	cactccact	3540

```

gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt 3600
ctgggggggtg ggggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat 3660
gctgggggatg cgggtgggctc tatgggtacc caggtgctga agaattgacc cggttcctcc 3720
tgggccagaa agaagcaggc acatccccct ctctgtgaca caccctgtcc acgccccctgg 3780
ttcttagttc cagccccact cataggacac tcatagctca ggagggtcc gccttcaatc 3840
ccacccgcta aagtacttgg agcggctctc ccctccctca tcagcccacc aaaccaaacc 3900
tagcctccaa gagtggaag aaattaaagc aagataggct attaagtga gagggagaga 3960
aaatgcctcc aacatgtgag gaagtaatga gagaaatcat agaattttaa ggccatgatt 4020
taaggccatc atggccttaa tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 4080
ttcggctgcg gcgagcggta tcagctcact caaaggcggc aatacggtta tccacagaat 4140
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 4200
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 4260
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 4320
cccctggaag ctccctcgct cgctctcctg ttccgacctt gccgcttacc ggataacctgt 4380
ccgcctttct cccttcggga agcgtggcgc tttctcatag ctacgctgt aggtatctca 4440
gttcggtgta ggtcgcttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg 4500
accgctgctc ctatccggt aactatcgtc ttgagtccaa cccggtaga cagacttat 4560
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 4620
cagagttctt gaagtggtag cctaactacg gctacactag aagaacagta tttggtatct 4680
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 4740
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 4800
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaaacgaa 4860
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 4920
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaaact tggctgaca 4980
gttaccaatg cttaatcagt gaggcacctt tctcagcgat ctgtctattt cgttcatcca 5040
tagttgcctg actcgggggg ggggggcgct gaggtctgcc tctgaagaa ggtgttgctg 5100
actcatacca ggcctgaatc gccccatcat ccagccagaa agtgaggag ccacggttga 5160
tgagagcttt gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac 5220
ggtctgctt gtcgggaaga tgcgtgatct gatccttcaa ctacgcaaaa gttcgattta 5280
ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt 5340
aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc 5400
aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc 5460
gaggcagttc cataggatgg caagatcctg gtatcgggtc gcgattccga ctgctccaac 5520
atcaatacaa cctattaatt tccccctgct aaaaataagg ttatcaagtg agaaatcacc 5580
atgagtgaag actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg 5640
ttcaacaggc cagccattac gctcgtcac aaaatcactc gcatcaacca aaccggtatt 5700
cattcgatgat tgcgcctgag cgagacgaaa tacgcatcg ctgttaaaag gacaattaca 5760
aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa tattttcacc 5820
tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg cagtggtag 5880
taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaattc 5940
cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc 6000
atgtttcaga aacaactctg gcgcacggg cttcccatac aatcgataga ttgtcgacc 6060
tgattgccc acattatcgc gagccattt ataccatat aaatcagcat ccatgttga 6120
atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa cacccttgt 6180
attactgttt atgtaagcag acagttttat tgttcatgat gatataattt tatcttgtgc 6240
aatgtaacat cagagatttt gagacacaac gtggctttcc ccccccccc attattgaag 6300
catttatcag ggttattgtc tcatgagcgg atacatatat gaatgtattt agaaaaataa 6360
acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaaccat 6420
tattatcatg acattaacct ataaaaatag gcgtatcacg aggcccttcc gtc 6473

```

&lt;210&gt; 101

&lt;211&gt; 6554

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 101

```

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgctt acataactta cggtaaattg 420
ccgcctggc tgaccgccc accgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgcat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcgat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca ccccttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttatgt accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat ggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggaacct gtgggtgacc gtgtactacg 2040
gcgtgcccg gtggaaggag gccaaagcca cctgttctg cgcagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgctg cgtgccacc gacccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340
gcgccatcac ccaggcctgc cccaagggtg gcttcgaccc catcccatc cactactgcy 2400
ccccgcgg ctacgccatc ctgaagtgc acaacaagac ctccaacggc accggcccc 2460
gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccggtgtg agcaccagc 2520
tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcacccgcc 2640
ccaacaacaa caccgcaag agcatccgca tcggccccgg ccaggccttc tacgccaccg 2700
gcgagatcat cggcgacatc cgccaggccc actgcaacat cagccgcacc gcctggaaca 2760
agaccctgca ggaggtgggc aagaagctgg ccgagcactt cccaacaag gccatcaagt 2820
tcgccaagca cagcggcggc gacctggaga tcaccacca cagcttcaac tgccgcggcg 2880
agttcttcta ctgcaacacc cgatcgatcc cctgccgcat caagcagatc atcaacatgt 2940

```



```

ggcagggcgt gggccgcgcc atgtacgccc ccccatcgga gggcatcctg acctgccgca 3000
gcaacatcac cggcctgctg ctgaccgcgc acggcggcac cggcatgcac gacaccgaga 3060
tcttccgccc cgagggcggc gacatgcgcg acaactggcg cagcgagctg tacaagtaca 3120
aggtggtgga gatcaagccc ctgggcatcg cccccaccaa ggccaagctt accgtgcagg 3180
tgcgccagct gctgagcggc atcgtgcagc agcagagcaa cctgctgcgc gccatcgagg 3240
cccagcagca catgctgcag ctgaccgtgt ggggcatcaa gcagctgcag acccgcgtag 3300
tggccatcga gcgctacctg cgcgaccagc agctcgagga catctggaac aacatgacct 3360
ggatgcagtg ggaccgcgag atcagcaact acaccaacac catctaccgc ctgctggagg 3420
acagccagaa ccagcaggag aagaacgagc aggacctgct ggccctggac aagtggcaga 3480
acctgtggac ctggttcggc atcaccaact ggctgtggtg aacacgtggg atccagatct 3540
gctgtgcctt ctagtgcga gccatctgtt gtttgcctt ccccgctgcc ttccttgacc 3600
ctggaagggt ccactccac tgtcctttcc taataaaatg aggaaattgc atcgcatgtt 3660
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 3720
tgggaagaca atagcaggca tgctggggat gcggtgggct ctatgggtac ccaggtgctg 3780
aagaattgac ccggttcctc ctgggccaga aagaagcagg cacatccctt tctctgtgac 3840
acaccctgtc caccgcccctg gttcttagtt ccagccccac tcataggaca ctcatagctc 3900
aggaggggtc cgccttcaat cccaccgct aaagtacttg gagcggtctc tccctccctc 3960
atcagcccac caaaccaaac ctagcctcca agagtgggaa gaaattaaag caagataggc 4020
tattaagtgc agagggagag aaaatgcctc caacatgtga ggaagtaatg agagaaatca 4080
tagaatttta aggccatgat ttaaggccat catggcctta atcttccgct tcctcgctca 4140
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 4200
taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 4260
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 4320
cccctgacga gcatcacaaa aatcgacgct caagtacagag gtggcgaaac ccgacaggac 4380
tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct 4440
tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata 4500
gctcacgctg taggtatctc agttcggtgt aggtcggtcg ctccaagctg ggctgtgtgc 4560
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 4620
accggttaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 4680
cgaggtatgt aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 4740
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 4800
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc 4860
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 4920
ctgacgctca gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa 4980
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 5040
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 5100
tctgtctatt tcgttcatcc atagtgcct gactcggggg gggggggcgc tgaggtctgc 5160
ctcgtgaaga aggtgttgct gactcatacc aggcctgaat cggccatca tccagccaga 5220
aagtgagggg gccacgggtg atgagagctt tggtgtaggt ggaccagttg gtgattttga 5280
acttttgctt tgccacggaa cggctctgct tgtcgggaag atgcgtgatc tgatccttca 5340
actcagcaaa agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct 5400
ctgccagtggt tacaaccaat taaccaatcc tgattagaaa aactcatcga gcatcaaatg 5460
aaactgcaat ttattcatac caggattatc aataccatat ttttgaaaaa gccgtttctg 5520
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcgggtc 5580
tgcgattccg actcgtccaa catcaataca acctattaat tccctcgt caaaaataag 5640
gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt 5700
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact 5760
cgcataacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgac 5820
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga aactgcccag 5880
cgcatacaaa atattttcac ctgaatcagg atatttctt aatacctgga atgctgtttt 5940
cccggggatc gcagtgggtg gtaaccatgc atcatcagga gtacggataa aatgcttgat 6000
ggtcgggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 6060
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcacggt gcttcccata 6120
caatcgatag attgtcgcac ctgattgcc gacattatcg cgagccatt tatacccata 6180
taaatacga tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat 6240

```



```

atggctcata acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 6300
tgatatatatt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 6360
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 6420
tgaatgtatt tagaaaaata aacaaatagg gggtccgcgc acatttcccc gaaaagtgcc 6480
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6540
gagggccttt cgtc
6554

```

&lt;210&gt; 102

&lt;211&gt; 6671

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 102

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggatgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gbatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgtagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcaccttg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgcagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgctg cgtgccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220

```

tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	atcgaagaac	tgcagcttca	2340
acatgaccac	cgaggtgcgc	gacaagcgcg	agaaggtgca	cgccctgttc	taccgcctgg	2400
acatcgtgcc	cctgaagaac	gagagcagca	acaccagcgg	cgactaccgc	ctgatcaact	2460
gcaacaccag	cgccatcacc	caggcctgcc	ccaaggtgag	cttcgacccc	atcccccattc	2520
actactgcgc	ccccgccggc	tacgccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	2580
ccggccccctg	caacaacgtg	agcaccatcc	agtgcaccca	cggcaccaag	cccgtggtga	2640
gcacccagct	gctgctgaac	ggcagcctgg	ccgaggagga	gatcatcatc	cgcagcaaga	2700
acctgaccga	caacgtgaag	accatcatcg	tgcacctgaa	cgagagcgtg	gagatcaact	2760
gcacccgccg	atcgccccac	tgcaacatca	gccgcaccgc	ctggaacaag	accctgcagg	2820
aggtgggcaa	gaagctggcc	gagcacttcc	ccaacaaggc	catcaagttc	gccaagcaca	2880
gcggcggcga	cctggagatc	accacccaca	gcttcaactg	ccgcggcgag	ttcttctact	2940
gcaacaccag	cagcctgttc	aacagcacct	acacccccaa	cagcaccgag	aacatcaccg	3000
gcaccgagaa	cagcatcatc	accatccccct	gccgcatcaa	gcagatcatc	aacatgtggc	3060
agggcgtggg	ccgcgccatg	tacgcccccc	ccatcgaggg	catcctgacc	tgccgcagca	3120
acatcaccgg	cctgctgctg	accgcgcagc	gcggcaccgg	catgcacgac	accgagatct	3180
tccgccccga	gggcggcgac	atgcgcgaca	actggcgag	cgagctgtac	aagtacaagg	3240
tgggtggagat	caagccccctg	ggcatcgccc	ccaccaaggc	caagcttacc	gtgcaggtgc	3300
gccagctgct	gagcggcatc	gtgcagcagc	agagcaacct	gctgcgcgcc	atcgaggccc	3360
agcagcacat	gctgcagctg	accgtgtggg	gcatcaagca	gctgcagacc	cgcgtgctgg	3420
ccatcgagcg	ctacctgcgc	gaccagcagc	tcgaggacat	ctggaacaac	atgacctgga	3480
tgcagtggga	ccgcgagatc	agcaactaca	ccaacaccat	ctaccgcctg	ctggaggaca	3540
gccagaacca	gcaggagaag	aacgagcagg	acctgctggc	cctggacaag	tggcagaacc	3600
tgtggacctg	gttcggcatc	accaactggc	tgtggtgaac	acgtgggac	cagatctgct	3660
gtgccttcta	gttgccagcc	atctgttgtt	tgccccctcc	ccgtgccttc	cttgaccctg	3720
gaaggtgcca	ctcccactgt	ccttttctaa	taaaatgagg	aaattgcac	gcattgtctg	3780
agtaggtgtc	attctattct	gggggggtggg	gtggggcagg	acagcaaggg	ggaggattgg	3840
gaagacaata	gcaggcatgc	tggggatgcg	gtgggtctta	tgggtacca	ggtgctgaag	3900
aattgacccg	gttcctcctg	ggccagaaaag	aagcaggcac	atccccctct	ctgtgacaca	3960
ccctgtccac	gcccctgggt	cttagttcca	gccccactca	taggacactc	atagctcagg	4020
agggctccgc	cttcaatccc	accgcctaaa	gtacttggag	cggtctctcc	ctccctcatc	4080
agcccaccaa	accaaacctc	gcctccaaga	gtgggaagaa	attaaagcaa	gataggctat	4140
taagtgcaga	gggagagaaa	atgcctccaa	catgtgagga	agtaatgaga	gaaatcatag	4200
aattttaagg	ccatgattta	aggccatcat	ggccttaatc	ttccgcttcc	tcgctcactg	4260
actcgtgcg	ctcggtcggt	cggtgcggc	gagcgggtatc	agctcactca	aaggcggtaa	4320
tacgggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	4380
aaaaggccag	gaaccgtaaa	aaggccgcgt	tgttggcgtt	tttccatagg	ctccgcccc	4440
ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaacccg	acaggactat	4500
aaagatacca	ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctggt	ccgaccctgc	4560
cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	4620
cacgctgtag	gtatctcagt	tcgggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	4680
aaccccccg	tcagcccagc	cgtgcgcct	tatccggtaa	ctatcgtctt	gagtccaacc	4740
cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	4800
ggtatgtagg	cgggtctaca	gagttcttga	agtgggtggc	taactacggc	tacactagaa	4860
gaacagtatt	tggatatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	4920
gctcttgatc	cggcaaacaa	accaccgctg	gtagcgggtg	tttttttgtt	tgcaagcagc	4980
agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acgggggtctg	5040
acgctcagtg	gaacgaaaac	tcacgttaag	ggattttggg	catgagatta	tcaaaaagga	5100
tcttcaccta	gatcctttta	aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	5160
agtaaacctg	gtctgacagt	taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	5220
gtctatcttg	ttcatccata	gttgccctgac	tcgggggggg	ggggcgctga	ggtctgcctc	5280
gtgaagaagg	tgttgctgac	tcataaccagg	cctgaatcgc	cccatcatcc	agccagaaag	5340
tgagggagcc	acggttgatg	agagctttgt	tgtaggtgga	ccagttgggtg	atthttgaact	5400
tttgctttgc	cacggaacgg	tctgcgttgt	cgggaagatg	cgtgatctga	tccttcaact	5460
cagcaaaaagt	tcgattttat	caacaaagcc	gccgtcccgt	caagtcagcg	taatgctctg	5520

```

ccagtgttac aaccaattaa ccaattctga ttagaaaaac tcatcgagca tcaaatgaaa 5580
ctgcaattta ttcatatcag gattatcaat accatatttt tgaaaaagcc gtttctgtaa 5640
tgaaggagaa aactcaccga ggcagttcca taggatggca agatcctggg atcgggtctgc 5700
gattccgact cgtccaacat caatacaacc tattaatttc ccctcgtcaa aaataagggt 5760
atcaagtgag aaatcaccat gagtgacgac tgaatccggg gagaatggca aaagcttatg 5820
catttctttc cagacttggt caacaggcca gccattacgc tcgtcatcaa aatcactcgc 5880
atcaaccaa ccgttattca ttcgtgattg cgcctgagcg agacgaaata cgcgatcgc 5940
gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc 6000
atcaacaata ttttcacctg aatcaggata ttcttctaata acctggaatg ctgttttccc 6060
ggggatcgca gtgggtgagta accatgcac atcaggagta cggataaaat gcttgatggg 6120
cggagaggc ataaattccg tcagccaggt tagtctgacc atctcatctg taacatcatt 6180
ggcaacgcta cctttgccat gtttcagaaa caactctggc gcacggggct tcccatacaa 6240
tcgatagatt gtcgcacctg attgcccagc attatcgca gcccatatat acccatataa 6300
atcagcatcc atgttggaat ttaatcgcg cctcgagcaa gacgtttccc gttgaatatg 6360
gctcataaca ccccttgat tactgtttat gtaagcagac agttttattg ttcattgatga 6420
tatattttta tcttggtgca tgtaacatca gagattttga gacacaacgt ggctttcccc 6480
cccccccat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 6540
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 6600
tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 6660
gccctttcgt c 6671

```

&lt;210&gt; 103

&lt;211&gt; 6602

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 103

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg acctatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380

```



```

cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggcttttcc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatt cagcgcgaact ggaagcagtg gtggatctgg ggcatcctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgaagaac tgcagcttca 2340
acatgaccac cgaggtgcgc gacaagcgcg agaaggtgca cgccctgttc taccgcctgg 2400
acatcgtgcc cctgaagaac gagagcagca acaccagcgg cgactaccgc ctgatcaact 2460
gcaacaccag cgccatcacc caggcctgcc ccaaggtgag cttcgacccc atccccatcc 2520
actactgcgc ccccgccggc tacgccatcc tgaagtgcaa caacaagacc ttcaacggca 2580
ccggcccctg caacaacgtg agcaccatcc agtgcaccca cggcaccaag cccgtggtga 2640
gcaccagct gctgctgaac ggcagcctgg ccgaggagga gatcatcacc cgcagcaaga 2700
acctgaccga caacgtgaag accatcatcg tgcacctgaa cgagagcgtg gagatcaact 2760
gcaccgccc atcgccccc tgcaacatca gccgcaccgc ctggaacaag accctgcagg 2820
aggtgggcaa gaagctggcc gagcacttcc ccaacaaggc catcaagttc gccaagcaca 2880
gcggcgggca cctggagatc accaccaca gcttcaactg ccgcgggcag ttcttctact 2940
gcaacacccg atcgatcccc tgccgcatca agcagatcat caacatgtgg cagggcgtgg 3000
gccgcgccat gtacgcccc cccatcgagg gcatcctgac ctgccgcagc aacatcaccg 3060
gcctgctgct gacccgcgac ggcggcaccg gcatgcacga caccgagatc ttccgccccg 3120
agggcgggca catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtggaga 3180
tcaagcccct gggcatcgcc cccaccaagg ccaagcttac cgtgcagggt cgccagctgc 3240
tgagcggcat cgtgcagcag cagagcaacc tgcctgcgcg ,catcgaggcc cagcagcaca 3300
tgcctgcagc gaccgtgtgg ggcattcaagc agctgcagac ccgctgctg gccatcgagc 3360
gctacctgcg cgaccagcag ctcgaggaca tctggaacaa catgacctgg atgcagtggg 3420
accgcgagat cagcaactac accaacacca tctaccgcct gctggaggac agccagaacc 3480
agcaggagaa gaacgagcag gacctgctgg ccctggacaa gtggcagaac ctgtggacct 3540
ggttcgggcat caccaactgg ctgtggtgaa cacgtgggat ccagatctgc tgtgccttct 3600
agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaagggtgc 3660
actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 3720
cattctatcc tgggggggtg ggtggggcag gacagcaagg gggaggattg ggaagacaat 3780
agcaggcatg ctggggatgc ggtgggctct atgggtaccc aggtgctgaa gaattgacct 3840
ggttcctcct gggccagaaa gaagcaggca catccccctc tctgtgacac accctgtcca 3900
cgcccctggg tcttagttcc agcccactc ataggacact catagctcag gagggtccg 3960
ccttcaatcc caccgctaa agtacttggg ggggtctctc cctccctcat cagcccacca 4020
aaccaaacct agcctccaag agtggggaaga aattaaagca agataggcta ttaagtgcag 4080
agggagagaa aatgcctcca acatgtgagg aagtaatgag agaaatcata gaattttaag 4140
gccatgattt aaggccatca tggccttaat cttccgcttc ctcgctcact gactcgctgc 4200
gctcggctcg tccgctgcgg cgagcggtat cagctcactc aaaggcggta atacggttat 4260
ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca 4320
ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc 4380
atcacaaaaa tcgacgctca agtcagaggt ggcgaaacct gacaggacta taaagatacc 4440
aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgacctg ccgcttaccg 4500
gatacctgct cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta 4560
ggtatctcag ttcgggtgtg gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg 4620
ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtagac 4680

```

```

acgacttata gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag 4740
gcggtgctac agagttcttg aagtgggtggc ctaactacgg ctacactaga agaacagtat 4800
ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggg agctcttgat 4860
ccggcaaaca aaccaccgct ggtagcgggt gtttttttgt ttgcaagcag cagattacgc 4920
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacgggggtct gacgctcagt 4980
ggaacgaaaa ctacagttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct 5040
agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt 5100
ggtctgacag ttaccaatgc ttaatcagtg aggacacctat ctacgcgacg tgtctatttc 5160
gttcatccat agttgcctga ctcggggggg gggggcgctg aggtctgcct cgtgaagaag 5220
gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgaggggagc 5280
cacggttgat gagagctttg ttgtaggtgg accagttggg gattttgaac ttttgctttg 5340
ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag 5400
ttcgatttat tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct gccagtgtta 5460
caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt 5520
attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga 5580
aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcggtctg cgattccgac 5640
tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt tatcaagtga 5700
gaaatcacca tgagtgcga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt 5760
ccagacttgt tcaacaggcc agccattacg ctcgatcatc aaatcactcg catcaaccaa 5820
accgttatcc attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg 5880
acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg catcaacaat 5940
atcttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc 6000
agtgggtgag aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg 6060
cataaattcc gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct 6120
acctttgcca tgtttcagaa acaactctgg cgcacggggc ttcccatata atcgatagat 6180
tgtcgccact gattgcccga cattatcgcg agcccattta taccatata aatcagcatc 6240
catgttgga tttaatcgcg gcctcgagca agacgtttcc cgttgaatat gggtcataac 6300
accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg atatatattt 6360
atcttggtgca atgtaacatc agagattttg agacacaacg tggctttccc cccccccca 6420
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatattg aatgtattta 6480
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta 6540
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg 6600
tc

```

&lt;210&gt; 104

&lt;211&gt; 6683

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 104

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttgggtcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaatagggg ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600

```



tgacggtaaa	tggccccgct	ggcattatgc	ccagttacatg	accttatggg	acttttcctac	660
ttggcagtac	atctacgtat	tagtcategc	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggatgt	ccaagtctcc	accccatgta	780
cgtcaatggg	agtttgtttt	ggcaccacaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcc	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttacaattt	cacatatata	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggagc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgtctcg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagc	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcgaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgac	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccg	gtggaaggag	gccaagacca	ccctgttctg	cgccagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	atcgaagaac	tgcagcttca	2340
acatgaccac	cgaggtgcgc	gacaagcgcg	agaagggtgca	cgccctgttc	taccgcctgg	2400
acatcgtgcc	cctgaagaac	gagagcagca	acaccagcgg	cgactaccgc	ctgatcaact	2460
gcaacaccag	cgccatcacc	caggcctgcc	ccaagggtgag	cttcgacccc	atccccatcc	2520
actactgcgc	ccccgcccgg	tacgccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	2580
ccggccccctg	caacaacgtg	agcaccatcc	agtgcaccca	cggcaccaag	cccgtggtga	2640
gcacccagct	gctgctgaac	ggcagcctgg	ccgaggagga	gatcatcatc	cgcagcaaga	2700
acctgaccga	caacgtgaag	accatcatcg	tgcacctgaa	cgagagcgtg	gagatcaact	2760
gcacccgccc	caacaacaac	acccgcaaga	gcatccgcat	cggccccggc	caggccttct	2820
acgccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacatc	agccgcaccg	2880
cctggaacaa	gaccctgcag	gaggtgggca	agaagctggc	cgagcacttc	cccaacaagg	2940
ccatcaagtt	cgccaagcac	agcggcgggc	acctggagat	caccacccac	agcttcaact	3000
gccgcggcga	gttctttctac	tgaacacccc	gatcgatccc	ctgccgcac	aagcagatca	3060
tcaacatgtg	gcaggggcgtg	ggccgcgcca	tgtacgcccc	ccccatcgag	ggcatcctga	3120
cctgccgcag	caacatcacc	ggcctgctgc	tgaccgcgca	cggcggcacc	ggcatgcacg	3180
acaccgagat	cttccgcccc	gagggcgggc	acatgcgcga	caactggcgc	agcgagctgt	3240
acaagtacaa	ggtggtggag	atcaagcccc	tgggcatcgc	ccccaccaag	gccaagctta	3300
ccgtgcaggt	gcgccagctg	ctgagcggca	tcgtgcagca	gcagagcaac	ctgctgcgcg	3360
ccatcgaggc	ccagcagcac	atgctgcagc	tgaccgtgtg	gggcatcaag	cagctgcaga	3420
cccgcgtgct	ggccatcgag	cgctacctgc	gcgaccagca	gctcgaggac	atctggaaca	3480
acatgacctg	gatgcagtgg	gaccgcgaga	tcagcaacta	caccaacacc	atctaccgcc	3540
tgctggagga	cagccagaac	cagcaggaga	agaacgagca	ggacctgctg	gccctggaca	3600
agtggcagaa	cctgtggacc	tggttcggca	tcaccaactg	gctgtggtga	acacgtggga	3660
tccagatctg	ctgtgccttc	tagttgccag	ccatctgttg	tttgcccttc	ccccgtgcct	3720
tccttgaccc	tgggaagggtg	cactcccact	gtcctttcct	aataaaaatga	ggaaattgca	3780
tcgcattgtc	tgagttaggtg	tcattctatt	ctgggggggtg	gggtggggca	ggacagcaag	3840
ggggaggatt	gggaagacaa	tagcaggcat	gctgggggatg	cgggtgggctc	tatgggtacc	3900

cagggtgctga agaattgacc cggttcctcc tggggccagaa agaagcaggc acatccccctt 3960  
ctctgtgaca caccctgtcc acgccccctgg ttcttagtcc cagccccact cataggacac 4020  
tcatagctca ggagggctcc gccttcaatc ccaccgccta aagtacttgg agcggctctt 4080  
ccctccctca tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4140  
aagataggct attaatgca gagggagaga aaatgcctcc aacatgtgag gaagtaatga 4200  
gagaaatcat agaattttta ggccatgatt taaggccatc atggccttaa tcttccgctt 4260  
cctcgctcac tgactcgctg cgctcggtcg ttcggtgctg gcgagcggta tcagctcact 4320  
caaaggcggc aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag 4380  
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 4440  
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 4500  
cgacaggact ataaagatac caggcggttc cccctggaag ctccctcgctg cgctctcctg 4560  
ttccgaccct gccgcttacc ggatacctgt ccgcttctct cccttcggga agcgtggcg 4620  
tttctcatag ctacgctgt aggtatctca gttcggtgta ggtcgctcgc tccaagctgg 4680  
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggc aactatcgtc 4740  
ttgagtccaa cccggtaaga cagcacttat cgccactggc agcagccact ggtaacagga 4800  
ttagcagagc gaggtatgta ggcgggtgta cagagttctt gaagtgggtg cctaactacg 4860  
gctacactag aagaacagta tttgggtatct gcgctctgct gaagccagtt accttcggaa 4920  
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggc ggtttttttg 4980  
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 5040  
ctacggggtc tgacgctcag tggaaacgaa actcacgtta agggattttg gtcattgagat 5100  
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 5160  
aaagtatata tgagtaaact tgggtctgaca gttaccaatg cttaatcagt gaggcaccta 5220  
tctcagcgat ctgtctatct cgttcatcca tagttgcctg actcgggggg gggggggcgt 5280  
gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc gccccatcat 5340  
ccagccagaa agtgaggag ccacgggtga tgagagcttt gttgtagggtg gaccagttgg 5400  
tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct 5460  
gatccttcaa ctacgcaaaa gttcgattta ttcaacaaag ccgcgctccc gtcaagtcag 5520  
cgtaatgctc tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag 5580  
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag 5640  
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5700  
gtatcggtct gcgattccga ctcgctccaa atcaatacaa cctattaatt tcccctcgtc 5760  
aaaaataagg ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg 5820  
caaaagctta tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc 5880  
aaaatcactc gcatcaacca aaccgttatt cattcgatg tgccctgag cgagacgaaa 5940  
tacgcgatcg ctgttaaaa gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa 6000  
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 6060  
tgctgttttc ccggggatcg cagtgggtgag taacctatgca tcatcaggag tacggataaa 6120  
atgcttgatg gtcggaagag gcataaatc cgtcagccag tttagtctga ccatctcatc 6180  
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga acaactctg gcgcacggg 6240  
cttcccatat aatcgataga ttgtcgccac tgattgccc acattatcgc gagccatttt 6300  
ataccatatt aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc 6360  
ccgttgaaata tggctcataa cacccttgtt attactgttt atgtaagcag acagttttat 6420  
tgttcatgat gatataatct tatcttgtgc aatgtaacat cagagatttt gagacacaac 6480  
gtggctttcc ccccccccc attattgaag catttatcag ggttattgtc tcatgagcgg 6540  
atacatatct gaatgtatct agaaaaataa acaaataggg gttccgcgca catttccccg 6600  
aaaagtgcc cctgacgtct aagaaaccat tattatcatg acattaacct ataaaaatag 6660  
gcgtatcacg aggccctttc gtc 6683

&lt;210&gt; 105

&lt;211&gt; 6605

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 105

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac ttcccaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatacgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcaccttgg 1980
gcttctggat ggtgatgac tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400
ccagcgccat caccagggc tgccccaagg tgagcttcga cccatcccc atccactact 2460
gcgccccgc cggctacgcc atcctgaagt gcaacaaca gaccttcaac ggcaccggcc 2520
cctgcaacaa cgtgagcacc atccagtgc cccacggcac caagcccgtg gtgagcacc 2580
agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc aagaacctga 2640
ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcaccc 2700
gccgatcggc ccactgcaac atcagccgca ccgctggaa caagaccctg caggaggtgg 2760
gcaagaagct ggccgagcac tcccccaaca aggccatcaa gttcgccaag cacagcggcg 2820
gcgacctgga gatcaccacc cacagcttca actgccgcgg cgagttcttc tactgcaaca 2880
ccagcagcct gttcaacagc acctacacc ccaacagcac cgagaacatc accggcaccg 2940
agaacagcat catcaccatc ccctgccgca tcaagcagat catcaacatg tggcagggcg 3000
tgggcgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3060

```



```

ccggcctgct gctgacccgc gacggcggca ccggcatgca cgacaccgag atcttccgcc 3120
ccgagggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aaggtggtgg 3180
agatcaagcc cctgggcatc gccccacca aggccaaagt taccgtgcag gtgcgccagc 3240
tgctgagcgg catcgtgcag cagcagagca acctgctgcg cgccatcgag gccagcagc 3300
acatgctgca gctgaccgtg tggggcatca agcagctgca gacccgcgtg ctggccatcg 3360
agcgctacct gcgcgaccag cagctcgagg acatctggaa caacatgacc tggatgcagt 3420
gggaccgcga gatcagcaac tacaccaaca ccatctaccg cctgctggag gacagccaga 3480
accagcagga gaagaacgag caggacctgc tggccctgga caagtggcag aacctgtgga 3540
cctggttcgg catcaccaac tggctgtggt gaacacgtgg gatccagatc tgctgtgcct 3600
tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt 3660
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg 3720
tgtcattcta ttctgggggg tgggggtggg caggacagca agggggagga ttgggaagac 3780
aatagcaggc atgctgggga tgcggtgggc tctatgggta cccaggtgct gaagaattga 3840
cccgttcct cctgggccag aaagaagcag gcacatcccc ttctctgtga cacaccctgt 3900
ccacgcccc tggttcttagt tccagcccca ctcataggac actcatagct caggagggtc 3960
ccgccttcaa tcccaccgc taaagtactt ggagcggctc ctccctccct catcagccca 4020
ccaaaccaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagt 4080
cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4140
aaggccatga ttttaaggcca tcatggcctt aatcttccgc ttctctgctc actgactcgc 4200
tgcgctcggc cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc 4260
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4320
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 4380
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4440
accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 4500
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4560
gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc 4620
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtta 4680
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4740
taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag 4800
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagbt ggtagctctt 4860
gatccggcaa acaaacacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 4920
cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 4980
agtggaaacga aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca 5040
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 5100
cttggctctga cagttacca tgcctaatca gtgaggcacc tatctcagcg atctgtctat 5160
ttcgttcatc catagttgcc tgactcgggg gggggggggc ctgaggtctg cctcgtgaag 5220
aaggtgttgc tgactcatc caggcctgaa tcgccccatc atccagccag aaagtgaggg 5280
agccacggtt gatgagagct ttggtgtagg tggaccagtt ggtgattttg aacttttgct 5340
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccctc aactcagcaa 5400
aagttcgatt tattcaacaa agcgcgcgtc ccgtcaagtc agcgtaatgc tctgccagt 5460
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5520
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5580
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggtatcggc ctgcgattcc 5640
gactcgacca acatcaatac aacctattaa ttccccctcg tcaaaaataa ggttatcaag 5700
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5760
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5820
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5880
aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg aacactgcca gcgcataaac 5940
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6000
cgcagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6060
aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6120
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6180
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6240
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6300
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6360

```

```

tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6420
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6480
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6540
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6600
tcgtc 6605

```

&lt;210&gt; 106

&lt;211&gt; 6605

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 106

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatc tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagt gtggatctgg ggcaccttg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccg gtggaaggag gccaaagacca ccctgttctg cgcagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280

```



agctgacccc	cctgtgcggtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	ccgatcgatc	aactgcaaca	2400
ccagcgccat	cacccaggcc	tgccccaagg	tgagcttcga	ccccatcccc	atccactact	2460
gcgccccgc	cggctacgcc	atcctgaagt	gcaacaacaa	gaccttcaac	ggcaccggcc	2520
cctgcaacaa	cgtgagcacc	atccagtgc	cccacggcac	caagcccgtg	gtgagcacc	2580
agctgctgct	gaacggcagc	ctggccgagg	aggagatcat	catccgcagc	aagaacctga	2640
ccgacaacgt	gaagaccatc	atcgtgcacc	tgaacgagag	cgtggagatc	aactgcaccc	2700
gccgatcggc	ccactgcaac	atcagccgca	ccgcctggaa	caagaccctg	caggaggtgg	2760
gcaagaagct	ggccgagcac	ttccccaaca	aggccatcaa	gttcgccaag	cacagcggcg	2820
gcgacctgga	gatcaccacc	cacagcttca	actgccgcgg	cgagttcttc	tactgcaaca	2880
ccagcagcct	gttcaacagc	acctacaccc	ccaacagcac	cgagaacatc	accggcaccg	2940
agaacagcat	catcaccatc	ccctgcccga	tcaagcagat	catcaacatg	tggcagggcg	3000
tgggcccgcg	catgtacgcc	ccccccatcg	agggcatcct	gacctgccgc	agcaacatca	3060
ccggcctgct	gctgaccgcg	gacggcggca	ccggcatgca	cgacaccgag	atcttccgcc	3120
ccgagggcg	cgacatgcgc	gacaactggc	gcagcgagct	gtacaagtac	aaggtggtgg	3180
agatcaagcc	cctgggcatc	gccccacca	aggccaagct	taccgtgcag	gtgcgccagc	3240
tgctgagcgg	catcgtgcag	cagcagagca	acctgctgcg	cgccatcgag	gcccagcagc	3300
acatgctgca	gctgaccgtg	tggggcatca	agcagctgca	gacccgcgtg	ctggccatcg	3360
agcgtacct	gcgcgaccag	cagctcgagg	acatctggaa	caacatgacc	tggatgcagt	3420
gggaccgcga	gatcagcaac	tacaccaaca	ccatctaccg	cctgctggag	gacagccaga	3480
accagcagga	gaagaacgag	caggacctgc	tggccctgga	caagtggcag	aacctgtgga	3540
cctgggttcgg	catcaccaac	tggctgtggt	gaacacgtgg	gatccagatc	tgctgtgcct	3600
tctagttgcc	agccatctgt	tgtttgcccc	tccccgtgc	cttccttgac	cctggaagg	3660
gccactccca	ctgtcctttc	ctaataaaat	gaggaaattg	catcgcattg	tctgagtagg	3720
tgtcattcta	ttctgggggg	tggggtgggg	caggacagca	agggggagga	ttgggaagac	3780
aatagcaggc	atgctgggga	tgcggtgggc	tctatgggta	cccaggtgct	gaagaattga	3840
cccggttcct	cctgggcccag	aaagaagcag	gcacatcccc	ttctctgtga	cacaccctgt	3900
ccacgccccct	ggttcttagt	tccagcccca	ctcataggac	actcatagct	caggagggct	3960
ccgccttcaa	tcccaccgcg	taaagtactt	ggagcggctt	ctccctccct	catcagccca	4020
ccaaaccaaa	cctagcctcc	aagagtggga	agaaattaaa	gcaagatagg	ctattaagtg	4080
cagagggaga	gaaaatgcct	ccaacatgtg	aggaagtaat	gagagaaatc	atagaatttt	4140
aaggccatga	tttaaggcca	tcatggcctt	aatcttcgcg	ttcctcgctc	actgactcgc	4200
tgcgctcgg	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacgg	4260
tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaagg	cagcaaaaagg	4320
ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	4380
agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccagacagga	ctataaagat	4440
accaggcgtt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	4500
ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	4560
gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	4620
ccgttcagcc	cgaccgctgc	gccttatecg	gtaactatcg	tcttgagtcc	aaccgggtaa	4680
gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	4740
taggcggtgc	tacagagttc	ttgaagtgg	ggcctaacta	cggctacact	agaagaacag	4800
tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	4860
gatccggcaa	acaaaccacc	gctggtagcg	gtgggttttt	tgtttgcaag	cagcagatta	4920
cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	4980
agtggaacga	aaactcacgt	taagggatct	tgggtcatgag	attatcaaaa	aggatcttca	5040
cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	5100
cttggtctga	cagttacca	tgcttaate	gtgaggcacc	tatctcagcg	atctgtctat	5160
ttcgttcatc	catagttgcc	tgactcgggg	ggggggggcg	ctgaggtctg	cctcgtgaag	5220
aaggtgttgc	tgactcatac	caggcctgaa	tcgccccatc	atccagccag	aaagtgaggg	5280
agccacggtt	gatgagagct	ttgttgtagg	tggaccagtt	gggtatcttg	aacttttgct	5340
ttgccacgga	acggtctgcg	ttgtcgggaa	gatgcgtgat	ctgatccttc	aactcagcaa	5400
aagttcgatt	tattcaacaa	agccgcgcgt	ccgtcaagtc	agcgtaatgc	tctgccagtg	5460
ttacaaccaa	ttaaccaatt	ctgattagaa	aaactcatcg	agcatcaaat	gaaactgcaa	5520
tttattcata	tcaggattat	caataccata	tttttgaaaa	agccgtttct	gtaatgaagg	5580

```

agaaaactca ccgaggcagt tccataggat ggcaagatcc tgggtatcggc ctgcgattcc 5640
gactcgteca acatcaatac aacctattaa tttcccctcg tcaaaaataa gggtatcaag 5700
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaagct tatgcatttc 5760
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5820
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5880
aggacaatta caaacaggaa tcgaatgcaa ccggcgagg aacactgcca gcgcatcaac 5940
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6000
cgcagtgggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6060
aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6120
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6180
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6240
atccatgttg gaatttaate gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6300
aacaccctt gtattactgt ttatgtaagc agacagtttt attgttcatt atgatattt 6360
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6420
ccattattga agcatttatc aggggttattg tctcatgagc ggatacatat ttgaatgtat 6480
ttagaaaaat aaacaaatag gggttccgcy cacatttccc cgaaaagtgc cacctgacgt 6540
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6600
tcgtc 6605

```

&lt;210&gt; 107

&lt;211&gt; 6536

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 107

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcy tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt battaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccb ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccctgccc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg ggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500

```

```

agcggctcat ggtcgctcgg cagctccttg ctectaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcattctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca cctgttctg cgccagcgac gccaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca ccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400
ccagcgccat caccaggcc tgcaccaagg tgagcttcca ccccatcccc atccactact 2460
gcgccccgcg cggctacgcc atcctgaagt gcaacaacaa gaccttcaac ggcaccggcc 2520
cctgcaacaa cgtgagcacc atccagtgc cccacggcac caagcccggtg gtgagcacc 2580
agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc aagaacctga 2640
ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcacc 2700
gccgatcggc ccactgcaac atcagccgca ccgcctggaa caagaccctg caggagggtg 2760
gcaagaagct ggccgagcac ttcccaacaa aggccatcaa gttcgccaag cacagcggcg 2820
gcgacctgga gatcaccacc cacagcttca actgccgcgg cgagttcttc tactgcaaca 2880
ccgatcgat cccctgccgc atcaagcaga tcatcaacat gtggcagggc gtgggccgcg 2940
ccatgtacgc cccccccatc gagggcatcc tgacctgccg cagcaacatc accggcctgc 3000
tgctgacccg cgacggcggc accggcatgc acgacaccga gatcttccgc cccgagggcg 3060
gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg gagatcaagc 3120
ccctgggcat cgccccacc aaggccaage ttacctgca ggtgcgccag ctgctgagcg 3180
gcacgtgca gcagcagagc aacctgctgc gcgccatcga ggccagcag cacatgctgc 3240
agctgaccgt gtggggcatc aagcagctgc agaccgcgt gctggccatc gagcgctacc 3300
tgcgcgacca gcagctcgag gacatctgga acaacatgac ctggatgcag tgggaccgcg 3360
agatcagcaa ctacaccaac accatctacc gcctgctgga ggacagccag aaccagcagg 3420
agaagaacga gcaggacctg ctggccctgg acaagtggca gaacctgtgg acctggttcg 3480
gcacaccaa ctggctgtgg tgaacacgtg ggatccagat ctgctgtgcc ttctagttgc 3540
cagccatctg ttgtttgccc ctcccccggt ccttccctga ccttgggaagg tgccactccc 3600
actgtccttt cctaataaaa tgaggaaatt gcacgcatt gtctgagtag gtgtcattct 3660
attctggggg gtgggggtgg gcaggacagc aagggggagg attgggaaga caatagcagg 3720
catgctgggg atgcggtggg ctctatgggt acccaggtgc tgaagaattg acccggttcc 3780
tcctgggcca gaaagaagca ggcacatccc cttctctgtg acacaccctg tccacgcccc 3840
tggttcttag ttccagcccc actcatagga cactcatagc tcaggagggc tccgccttca 3900
atcccacccg ctaaagtact tggagcggtc tctccctccc tcatcagccc accaaacca 3960
acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag 4020
agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg 4080
atttaaggcc atcatggcct taatcttccg ctctctcgct cactgactcg ctgcgctcgg 4140
tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 4200
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 4260
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 4320
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 4380
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 4440
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 4500
tcagttcggg ttaggtcggt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc 4560
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta agacacgact 4620
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcggtg 4680
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggt 4740
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 4800

```



```

aacaaaccac cgctggtagc ggtgggttttt ttgtttgcaa gcagcagatt acgcgcagaa 4860
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 4920
aaaactcacg ttaagggatt ttgggtcatga gattatcaaa aaggatcttc acctagatcc 4980
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 5040
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 5100
ccatagttgc ctgactcggg gggggggggg gctgaggtct gcctcgtgaa gaagggtgtg 5160
ctgactcata ccaggcctga atcgcccat catccagcca gaaagtgagg gagccacggt 5220
tgatgagagc tttgtttagt gtggaccagt tgggtatttt gaacttttgc tttgccacgg 5280
aacggtctgc gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat 5340
ttattcaaca aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca 5400
attaaccaat tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat 5460
atcaggatta tcaataccat atttttgaaa aagccgttct tgtaatgaag gagaaaactc 5520
accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc 5580
aacatcaata caacctatta atttccctc gtcaaaaata aggttatcaa gtgagaaatc 5640
accatgagtg acgactgaat ccggtgagaa tggcaaaaagc ttatgcattt ctttccagac 5700
ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctgcgatcaa ccaaaccgtt 5760
attcatctgt gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt 5820
acaaacagga atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatatttct 5880
acctgaatca ggatattctt ctaataacctg gaatgctgtt tccccgggga tcgcagtggg 5940
gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa 6000
ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt 6060
gccatgttct agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc 6120
acctgattgc ccgacattat cgcgagccca tttataccca tataaatcag catccatgtt 6180
ggaatttaat cgcggcctcg agcaagacgt ttcccggtga atatggctca taacaccctt 6240
tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg 6300
tgcaatgtaa catcagagat tttgagacac aacgtggctt tccccccccc cccattattg 6360
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6420
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6480
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6536

```

&lt;210&gt; 108

&lt;211&gt; 6617

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 108

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```

ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttaciaaatt	cacatatata	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcgagagt	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcccttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgac	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaagacca	ccctgttctg	cgcacgcgac	gccaaggcct	2100
acgacgccga	ggtgcacaa	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	ccgatcgatc	aactgcaaca	2400
ccagcgccat	cacccaggcc	tgccccaaag	tgagcttcga	ccccatcccc	atccactact	2460
gcgcccccg	cggctacgcc	atcctgaagt	gcaacaacaa	gaccttcaac	ggcaccggcc	2520
cctgcaacaa	cgtgagcacc	atccagtgca	cccacggcac	caagcccgtg	gtgagcacc	2580
agctgctgct	gaacggcagc	ctggccgagg	aggagatcat	catccgcagc	aagaacctga	2640
ccgacaacgt	gaagaccatc	atcgtgcacc	tgaacgagag	cgtggagatc	aactgcacc	2700
gccccaaaca	caacacccgc	aagagcatcc	gcacggccc	cggccaggcc	ttctacgcca	2760
ccggcgagat	catcggcgac	atccgccagg	cccactgcaa	catcagccgc	accgcctgga	2820
acaagacctt	gcaggagggtg	ggcaagaagc	tggccgagca	cttccccaac	aaggccatca	2880
agttcgccaa	gcacagcggc	ggcgacctgg	agatcaccac	ccacagcttc	aactgccgcg	2940
gcgagttctt	ctactgcaac	acccgatcga	tccccgtccg	catcaagcag	atcatcaaca	3000
tgtggcaggg	cgtgggcggc	gccatgtacg	ccccccccat	cgagggcatc	ctgacctgcc	3060
gcagcaacat	caccggcctg	ctgctgaccc	gcgacggcgg	caccggcatg	cacgacaccg	3120
agatcttccg	ccccgagggc	ggcgacatgc	gcgacaactg	gcgcagcgag	ctgtacaagt	3180
acaagggtgg	ggagatcaag	cccctgggca	tcgccccac	caaggccaag	cttaccgtgc	3240
aggtgcgcca	gctgctgagc	ggcatcgtgc	agcagcagag	caacctgctg	cgcgccatcg	3300
aggcccagca	gcacatgctg	cagctgaccg	tgtggggcat	caagcagctg	cagacccgcg	3360
tgctggccat	cgagcgctac	ctgcgcgacc	agcagctcga	ggacatctgg	aacaacatga	3420
cctggatgca	gtgggaccgc	gagatcagca	actacaccaa	caccatctac	cgcctgctgg	3480
aggacagcca	gaaccagcag	gagaagaacg	agcaggacct	gctggccctg	gacaagtggc	3540
agaacctgtg	gacctgggtc	ggcatcacca	actggctgtg	gtgaacacgt	gggatccaga	3600
tctgctgtgc	cttctagttg	ccagccatct	gttgtttgcc	cctcccccg	gccttccctg	3660
accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcacgcat	3720
tgtctgagta	ggtgtcattc	tattctgggg	ggtgggggtg	ggcaggacag	caagggggag	3780
gattgggaag	acaatagcag	gcatgctggg	gatgcgggtg	gctctatggg	taccaggtg	3840
ctgaagaatt	gacccgggtc	ctcctgggcc	agaaagaagc	aggcacatcc	ccttctctgt	3900
gacacacctt	gtccacgccc	ctggttctta	gttccagccc	cactcatagg	acactcatag	3960
ctcaggaggg	ctccgccttc	aatcccaccc	gctaaagtac	ttggagcggg	ctctccctcc	4020
ctcatcagcc	caccaaacca	aacctagcct	ccaagagtgg	gaagaaatta	aagcaagata	4080
ggctattaag	tgcagaggga	gagaaaatgc	ctccaacatg	tgagggaagta	atgagagaaa	4140



tcatagaatt ttaaggccat gatttaagggc catcatggcc ttaatcttcc gcttcctcgc 4200  
tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4260  
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 4320  
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttget ggcgtttttc cataggctcc 4380  
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4440  
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4500  
ccctgccgct taccggatac ctgtccgcct ttctcccttc ggggaagcgtg gcgctttctc 4560  
atagctcacg ctgtaggtat ctcagttcgg ttaggtcgt tgcctccaag ctgggctgtg 4620  
tgacgaacc ccccgttcag cccgaccgct gcgccttate cggtaactat cgtcttgagt 4680  
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4740  
gagcgaggta ttagggcggc gctacagagt tcttgaagt gtggcctaac tacggctaca 4800  
ctagaagaac agtatttggc atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4860  
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca 4920  
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 4980  
ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatt agattatcaa 5040  
aaaggatctt cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta 5100  
tatatgagta aacttggctt gacagttacc aatgcttaat cagtgaggca cctatctcag 5160  
cgatctgtct atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc 5220  
tgctcgtga agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc 5280  
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5340  
tgaacttttg ctttgccacg gaacgggtctg cgttgctcggg aagatgcgtg atctgatect 5400  
tcaactcagc aaaagtctga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat 5460  
gctctgccag tggtacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5520  
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5580  
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5640  
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttccccct cgtcaaaaat 5700  
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag 5760  
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 5820  
actcgcatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 5880  
atcgctgtta aaaggacaat tacaacacag aatcgaatgc aaccggcgca ggaacactgc 5940  
cagcgcatca acaatatattt cacctgaatc aggatattct tctaatacct ggaatgctgt 6000  
tttcccgggg atcgcagtg tgagtaacca tgcacatca ggagtacgga taaaatgctt 6060  
gatggtcggg agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6120  
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6180  
atacaatcga tagattgtcg cacctgattg cccgacatta tccgcgagccc atttataccc 6240  
atataaatca gcatccatgt tggaaattta tgcgcgcctc gagcaagacg tttcccgttg 6300  
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6360  
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6420  
ttcccccccc cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6480  
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6540  
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6600  
cacgaggccc tttcgtc 6617

&lt;210&gt; 109

&lt;211&gt; 6713

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 109

tcgcgcgctt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60

```

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaag 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagecgaact ggaagcagtg gtggatctgg ggcatectgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccg gtggaaggag gccaaagacca cctgtttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggt caccgccctgt tctaccgctt ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520
cccaggcctg ccccaagggt agcttcgacc ccatecccat ccactactgc gccccgcgcg 2580
gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc caccggacca agcccggtgg gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcaccgc cgatcgcccc 2820
actgcaacat cagccgcacc gcctggaaca agaccctgca ggaggtgggc aagaagctgg 2880
ccgagcactt ccccaacaag gccatcaagt tcgccaagca cagcggcggc gacctggaga 2940
tcaccaccca cagcttcaac tgccgcggcg agttcttcta ctgcaacacc agcagcctgt 3000
tcaacagcac ctacaccccc aacagcaccg agaacatcac cggcaccgag aacagcatca 3060
tcaccatccc ctgccgcac aagcagatca tcaacatgtg gcagggcggt ggccgcgcca 3120
tgtacgcccc ccccatcgag ggcatectga cctgccgcag caacatcacc ggctgctgc 3180
tgaccgcga cggcggcacc ggcagtcacg acaccgagat ctcccgcccc gagggcggcg 3240
acatgcgcga caactggcgc agcgagctgt acaagtacaa ggtggtggag atcaagcccc 3300
tgggcatcgc cccaccaag gccaaagctt ccgtgcaggt gcgccagctg ctgagcggca 3360

```

tcgtgcagca	gcagagcaac	ctgctgcgcg	ccatcgaggc	ccagcagcac	atgctgcagc	3420
tgaccgtgtg	gggcatcaag	cagctgcaga	cccgcgtgct	ggccatcgag	cgctacctgc	3480
gcgaccagca	gctcgaggac	atctggaaca	acatgacctg	gatgcagtgg	gaccgcgaga	3540
tcagcaacta	caccaacacc	atctaccgcc	tgctggagga	cagccagaac	cagcaggaga	3600
agaacgagca	ggacctgctg	gccctggaca	agtggcagaa	cctgtggacc	tggttcggca	3660
tcaccaactg	gctgtggtga	acacgtggga	tccagatctg	ctgtgccttc	tagttgccag	3720
ccatctgttg	tttgcccctc	ccccgtgcct	tccttgaccc	tggaaagggtgc	cactcccact	3780
gtcctttcct	aataaaatga	ggaaattgca	tcgcattgtc	tgagtaggtg	tcattctatt	3840
ctgggggggtg	gggtggggca	ggacagcaag	ggggaggatt	gggaagacaa	tagcaggcat	3900
gctgggggatg	cggtggggctc	tatgggtacc	caggtgctga	agaattgacc	cggttcctcc	3960
tgggcccagaa	agaagcaggc	acatcccctt	ctctgtgaca	caccctgtcc	acgcccctgg	4020
ttcttagttc	cagccccact	cataggacac	tcatagctca	ggagggtctc	gccttcaatc	4080
ccacccgcta	aagtacttgg	agcggctctc	ccctcccctca	tcagcccacc	aaaccaaacc	4140
tagcctccaa	gagtgggaag	aaattaaagc	aagataggct	attaagtgca	gaggagaga	4200
aaatgcctcc	aacatgtgag	gaagtaatga	gagaaatcat	agaattttaa	ggccatgatt	4260
taaggccatc	atggccttaa	tcttcgcgtt	cctcgctcac	tgactcgctg	cgctcggctc	4320
ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggg	aatacggtta	tccacagaat	4380
caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	4440
aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	catcacaaaa	4500
atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	4560
cccctggaag	ctccctcgctg	cgctctcctg	ttccgacctc	gccgcttacc	ggatacctgt	4620
ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	4680
gttcgggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccg	4740
accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	4800
cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	ggcgggtgcta	4860
cagagtctct	gaagtgggtg	cctaactacg	gctacactag	aagaacagta	tttggtatct	4920
gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	4980
aaaccaccgc	tggtagcggg	ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	5040
aaggatctca	agaagatcct	ttgatctttt	ctacgggggc	tgacgctcag	tggaaacgaa	5100
actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	5160
taaattaaaa	atgaagtttt	aaatcaatct	aaagtatata	tgagtaaact	tggctctgaca	5220
gttaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	ctgtctatct	cgttcatcca	5280
tagttgcctg	actcgggggg	ggggggcgct	gaggtctgcc	tcgtgaagaa	ggtgttgctg	5340
actcatacca	ggcctgaatc	gccccatcat	ccagccagaa	agtgaggag	ccacgggtga	5400
tgagagcttt	gttgtaggtg	gaccagttgg	tgattttgaa	cttttgcttt	gccacggaac	5460
ggtctgcgtt	gtcgggaaga	tgcgtgatct	gatccttcaa	ctcagcaaaa	gttcgattta	5520
ttcaacaaag	ccgccgtccc	gtcaagtcag	cgtaatgctc	tgccagtgtt	acaaccaatt	5580
aaccaattct	gattagaaaa	actcatcgag	catcaaatga	aactgcaatt	tattcatatc	5640
aggattatca	ataccatatt	tttgaaaaag	ccgtttctgt	aatgaaggag	aaaactcacc	5700
gaggcagttc	cataggatgg	caagatcctg	gtatcgggtc	gcgattccga	ctcgtccaac	5760
atcaatacaa	cctattaatt	ttccctcgtc	aaaaataagg	ttatcaagtg	agaaatcacc	5820
atgagtgcg	actgaatccg	gtgagaatgg	caaaagctta	tgcattttct	tccagacttg	5880
ttcaacaggc	cagccattac	gctcgtcatc	aaaatcactc	gcatcaacca	aaccgttatt	5940
cattcgtgat	tgcgcctgag	cgagacgaaa	tacgcgatcg	ctgttaaaag	gacaattaca	6000
aacaggaatc	gaatgcaacc	ggcgcaggaa	cactgccagc	gcatcaacaa	tattttcacc	6060
tgaatcagga	tattcttcta	atacctggaa	tgctgttttc	ccggggatcg	cagtgggtgag	6120
taaccatgca	tcatcaggag	tacggataaa	atgcttgatg	gtcgggaagag	gcataaatcc	6180
cgtcagccag	tttagtctga	ccatctcatc	tgtaacatca	ttggcaacgc	tacctttgcc	6240
atgtttcaga	aacaactctg	gcgcacgggg	cttcccatac	aatcgataga	ttgtcgcacc	6300
tgattgccc	acattatcgc	gagcccattt	atacccatat	aaatcagcat	ccatgttgga	6360
atttaatcgc	ggcctcgagc	aagacgtttc	ccgttgaaata	tggctcataa	caccccttgt	6420
attactgttt	atgtaagcag	acagttttat	tgttcatgat	gatataatct	tatcttgtgc	6480
aatgtaacat	cagagatttt	gagacacaac	gtggctttcc	ccccccccc	attattgaag	6540
catttatcag	ggttattgtc	tcattgagcgg	atacatattt	gaatgtattt	agaaaaataa	6600
acaaataggg	gttccgcgca	catttccccg	aaaagtgcc	cctgacgtct	aagaaaccat	6660



tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtc

6713

&lt;210&gt; 110

&lt;211&gt; 6644

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 110

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcggtt acataactta cggtaaattg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttccctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcggtta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgcgtgc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacccctg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca cctgttcttg cgccagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggt cagccctgtt tctaccgctt ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgcctca 2520

```

```

cccaggcctg ccccaagggtg agcttcgacc ccatccccc at ccactactgc gccccgcgcg 2580
gctacgccat cctgaagtgc aacaacaaga ctttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc cacggcacca agcccgtggg gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcacccgc cgatcgcccc 2820
actgcaacat cagccgcacc gcctggaaca agaccctgca ggagggtggg aagaagctgg 2880
ccgagcactt ccccaacaag gccatcaagt tcgccaagca cagcggcggc gacctggaga 2940
tcaccacca cagcttcaac tgccgcggcg agttcttcta ctgcaacacc cgatcgatcc 3000
cctgcgcgat caagcagatc atcaacatgt ggcagggcgt gggccgcgc atgtacgcc 3060
cccccatcga gggcatcctg acctgccga gcaacatcac cggcctgctg ctgacccgcg 3120
acggcggcac cggcatgcac gacaccgaga tcttccgccc cgagggcggc gacatgcgcg 3180
acaactggcg cagcgagctg tacaagtaca aggtgggtga gatcaagccc ctgggcatcg 3240
ccccaccaaa ggccaagctt accgtgcagg tgcgccagct gctgagcggc atcgtgcagc 3300
agcagagcaa cctgctgcgc gccatcgagg ccagcagca catgctgcag ctgaccgtgt 3360
ggggcatcaa gcagctgcag acccgcgctg tggccatcga gcgctacctg cgcgaccagc 3420
agctcgagga catctggaac aacatgacct ggatgcagtg ggaccgcgag atcagcaact 3480
acaccaacac catctaccgc ctgctggagg acagccagaa ccagcaggag aagaacgagc 3540
aggacctgct ggcctggac aagtggcaga acctgtggac ctggttcggc atcaccaact 3600
ggctgtggtg aacacgtggg atccagatct gctgtgcctt ctagtgtcca gccatctgtt 3660
gtttgcccc ccccgctgcc ttccttgacc ctggaagggt ccactccac tgcctttcc 3720
taataaaatg aggaaattgc atcgcatgt ctgagtaggt gtcattctat tctggggggt 3780
ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat 3840
gcggtgggct ctatgggtac ccaggtgctg aagaattgac ccggttcctc ctgggccaga 3900
aagaagcagg cacatccctt tctctgtgac acaccctgtc cagcccttg gttcttagtt 3960
ccagccccac tcataggaca ctcatagctc aggagggtc cgccttcaat cccaccgct 4020
aaagtacttg gagcggcttc tccctccctc atcagccac caaaccaaac ctagcctcca 4080
agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc 4140
caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatgat ttaaggccat 4200
catggcctta atcttccgct tctcgcctca ctgactcgct gcgctcggtc gttcggtgc 4260
ggcgagcggg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 4320
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 4380
cgttgctggc gtttttccat aggtccgc cccctgacga gcacacaaa aatcgacgct 4440
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 4500
gctccctcgt gcgctctcct gttccgacct tgccgcttac cggatacctg tccgcctttc 4560
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 4620
aggtcgctcg ctccaagctg ggctgtgtgc acgaacccc cgttcagccc gaccgctgcg 4680
ccttatccgg taactatcgt cttgagtcca acccggtaa acacgactta tcgccactgg 4740
cagcagccac tggtaacagg attagcagag cgaggatgt aggcggtgct acagagttct 4800
tgaagtgggt gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 4860
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 4920
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 4980
aagaagatcc tttgatcttt tctacggggt ctgacgtca gtggaacgaa aactcacggt 5040
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 5100
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 5160
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 5220
gactcggggg gggggggcgc tgaggctctgc ctctgaaga aggtgttgct gactcatacc 5280
aggcctgaat cggcccatca tccagccaga aagtgaggga gccacggttg atgagagctt 5340
tggtgtaggt ggaccagttg gtgattttga acttttgctt tgccacggaa cggctctgcg 5400
tgtcgggaag atgcgtgatc tgatccttca actcagcaaa agttcgattt attcaacaaa 5460
gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtg tacaaccaat taaccaattc 5520
tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatal caggattatc 5580
aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac cgaggcagtt 5640
ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtecaa catcaatata 5700
acctattaat ttccctcgt caaaaataag gttatcaagt gagaaatcac catgagtgc 5760
gactgaatcc ggtgagaatg gcaaaagctt atgcatttct ttccagactt gttcaacagg 5820

```



```

ccagccatta cgctcgtcat caaaatcact cgcacaaacc aaaccggttat tcattcgtga 5880
ttgcgcctga gcgagacgaa atacgcgacg gctgttaaaa ggacaattac aaacaggaat 5940
cgaatgcaac cggcgcagga acactgccag cgcacaaaca atattttcac ctgaatcagg 6000
atattcttct aatacctgga atgctgtttt cccggggatc gcagtgggtga gtaaccatgc 6060
atcatcagga gtacggataa aatgcttgat ggtcggaaga ggcataaatt ccgtcagcca 6120
gttttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc catgtttcag 6180
aaacaactct ggcgcacatcg gcttcccata caatcgatag attgtcgcac ctgattgccc 6240
gacattatcg cgagcccatt tatacccata taaatcagca tccatgtttg aatttaatcg 6300
cggcctcagc caagacgttt cccgttgaat atggctcata acaccccttg tattactgtt 6360
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 6420
tcagagattt tgagacacaa cgtggctttc ccccccccc cattattgaa gcatttatca 6480
gggttattgt ctcacgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 6540
ggttccgcgc acatttcccc gaaaagtgc acctgacgtc taagaaacca ttattatcat 6600
gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtc 6644

```

&lt;210&gt; 111

&lt;211&gt; 6725

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 111

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgcctg cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgcctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740

```

```

gttgccggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgccgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatctcgcg cgcctctaga gatctcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacccctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccggt gtggaaggag gccaaagacca ccctgttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggt cagccctgt tctaccgcct ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520
cccaggcctg ccccaagggt agcttcgacc ccatcccat ccaactactg gcccccgccg 2580
gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc cagggacca agcccggtgt gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcaccgc cccaacaaca 2820
acaccgcaa gagcatccgc atcgccccg gccaggcctt ctacgccacc ggcgagatca 2880
tcggcgacat ccgccaggcc cactgcaaca tcagccgcac cgcttgaac aagacctgc 2940
aggaggtggg caagaagctg gccgagcact tcccaacaa ggccatcaag ttcgccaagc 3000
acagcggcgg cgacctggag atcaccaccc acagcttcaa ctgcccgggc gagttcttct 3060
actgcaacac ccgatcgatc ccctgccgca tcaagcagat catcaacatg tggcagggcg 3120
tgggcgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3180
ccggcctgct gctgaccgc gacggcggca ccggcatgca cgacaccgag atcttccgc 3240
ccgagggcgg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aagggtggtg 3300
agatcaagcc cctgggcatc gccccacca agccaagct taccgtgcag gtgcgccagc 3360
tgctgagcgg catcgtgcag cagcagagca acctgctgcg cgccatcgag gccagcagc 3420
acatgctgca gctgaccgtg tggggcatca agcagctgca gaccgcgtg ctggccatcg 3480
agcgtacct gcgcgaccag cagctcgagg acatctggaa caacatgacc tggatgcagt 3540
gggaccgcga gatcagcaac tacaccaaca ccatctaccg cctgctggag gacagccaga 3600
accagcagga gaagaacgag caggacctgc tggccctgga caagtggcag aacctgtgga 3660
cctggttcgg catcaccaac tggctgtggt gaacacgtgg gatccagatc tgctgtgcct 3720
tctagttgcc agccatctgt tgtttgcccc tcccccgctg cttccttgac cctggaagg 3780
gccactccca ctgtccttct ctaataaaat gaggaaattg catcgcatg tctgagtagg 3840
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 3900
aatagcagga atgctgggga tgcggtgggc tctatgggta ccaggtgct gaagaattga 3960
cccggttcct cctgggccag aaagaagcag gcacatcccc ttctctgtga cacacctgt 4020
ccacgcccct ggttcttagt tccagcccca ctcataggac actcatagct caggagggt 4080
ccgccttcaa tcccaccgc taaagtactt ggagcggctt ctccctccct catcagccca 4140
ccaaaccaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagt 4200
cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4260
aaggccatga ttaaggcca tcatggcctt aatcttccgc ttcctcgctc actgactgc 4320
tgcgtcggt cgttcggtg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg 4380
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4440
ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 4500
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4560
accaggcgtt tccccctgga agctccctcg tgcgtctcc tgttccgacc ctgcccgtta 4620
ccgatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4680
gtaggtatct cagttcgggt taggtcggtt gctccaagct gggctgtgtg cacgaacccc 4740
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa 4800
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4860
taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag 4920
tatgtggtat ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 4980
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 5040

```

```

cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5100
agtggaacga aaactcacgt taagggatct tgggtcatgag attatcaaaa aggatcttca 5160
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtata 5220
cttgggtctga cagttaccaa tgcttaatac gtgaggcacc tatctcagcg atctgtctat 5280
ttcgttcatc catagttgcc tgactcgggg gggggggggcg ctgaggtctg cctcgtgaag 5340
aaggtgttgc tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5400
agccacgggt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5460
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgaccttc aactcagcaa 5520
aagttcgatt tattcaacaa agccgccgct ccgtcaagtc agcgtaatgc tctgccagtg 5580
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5640
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5700
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggatcgggt ctgcgattcc 5760
gactcgtcca acatcaatac aacctattaa tttcccctcg tcaaaaataa ggttatcaag 5820
tgagaaatca ccatgagtga cgactgaatc cgggtgagaat ggcaaaagct tatgcatttc 5880
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5940
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 6000
aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg aacactgcca gcgcatacaac 6060
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6120
cgcagtgggt agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6180
aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6240
gctacctttg ccatgtttca gaaacaactc tggcgcacat ggcttcccat acaatcgata 6300
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataatcagc 6360
atccatgttg gaatttaate gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6420
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatataat 6480
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6540
ccattattga agcatttate agggttattg tctcatgagc ggatacataa ttgaatgtat 6600
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6660
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6720
tcgtc 6725

```

&lt;210&gt; 112

&lt;211&gt; 6824

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 112

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```



ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	btgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtggggtc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggtgct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggtcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcac	cagcgcaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgatc	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaaagacca	ccctgttctg	cgccagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtggggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	atcgaagaac	tgagcttca	2340
acatgaccac	cgaggtgcgc	gacaagcgcg	agaaggtgca	cgccctgttc	taccgcctgg	2400
acatcgtgcc	cctgaagaac	gagagcagca	acaccagcgg	cgactaccgc	ctgatcaact	2460
gcaacaccag	cgccatcacc	caggcctgcc	ccaaggtgag	cttcgacccc	atccccatcc	2520
actactgcgc	ccccgcgggc	tacgccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	2580
ccggccccctg	caacaacgtg	agcaccatcc	agtgcaccca	cggcaccaag	cccgtggtga	2640
gcacccagct	gctgctgaac	ggcagcctgg	ccgaggagga	gatcatcatc	cgcagcaaga	2700
acctgaccga	caacgtgaag	accatcatcg	tgcacctgaa	cgagagcgtg	gagatcaact	2760
gcacccgccc	caacaacaac	acccgcaaga	gcatccgcac	cggccccggc	caggccttct	2820
acgccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacatc	agccgcaccg	2880
cctggaacaa	gaccctgcag	gaggtgggca	agaagctggc	cgagcacttc	cccaacaagg	2940
ccatcaagtt	cgccaagcac	agcggcgggc	acctggagat	caccacccac	agcttcaact	3000
gccgcggcga	gttctttctac	tgcaacacca	gcagcctgtt	caacagcacc	tacaccccca	3060
acagcaccga	gaacatcacc	ggcaccgaga	acagcatcat	caccatcccc	tgccgcatca	3120
agcagatcat	caacatgtgg	cagggcggtg	gccgcgccat	gtacgcccc	ccatcgagg	3180
gcacccctgac	ctgccgcagc	aacatcaccg	gcctgctgct	gacccgcgac	ggcggcaccg	3240
gcacgcacga	caccgagatc	ttccgccccg	agggcgggcga	catgcgcgac	aactggcgca	3300
gcgagctgta	caagtacaag	gtgggtggaga	tcaagccctt	gggcatcgcc	cccaccaagg	3360
ccaagcttac	cgtgcaggtg	cgcagctgc	tgagcggcat	cgtgcagcag	cagagcaacc	3420
tgctgcgcgc	catcgaggcc	cagcagcaca	tgctgcagct	gaccgtgtgg	ggcatcaagc	3480
agctgcagac	ccgcgtgctg	gccatcgagc	gctacctgcg	cgaccagcag	ctcgaggaca	3540
tctggaacaa	catgacctgg	atgcagtggg	accgcgagat	cagcaactac	accaacacca	3600
tctaccgcct	gctggaggac	agccagaacc	agcaggagaa	gaacgagcag	gacctgctgg	3660
ccctggacaa	gtggcagaac	ctgtggacct	ggttcggcat	caccaactgg	ctgtggtaca	3720
tcaagatctt	catcaagatc	gtgggcggcc	tgatcggcct	gcgcatcatc	ttcgccgtgc	3780
tgagcatctg	aacacgtggg	atccagatct	gctgtgcctt	ctagttgcca	gccatctgtt	3840
gtttgcccc	cccccggtgc	ttccttgacc	ctggaagggtg	ccactcccac	tgctctttcc	3900
taataaaatg	aggaaattgc	atcgcatgtt	ctgagtaggt	gtcattctat	tctggggggg	3960
gggggtggggc	aggacagcaa	gggggaggat	tgggaagaca	atagcaggca	tgctggggat	4020
gcggtgggct	ctatgggtac	ccaggtgctg	aagaattgac	ccggttcctc	ctggggcaga	4080
aagaagcagg	cacatccctt	tctctgtgac	acaccctgtc	cacgcccctg	gttcttagtt	4140

```

ccagccccac tcataggaca ctcatagctc aggagggctc cgccttcaat cccacccgct 4200
aaagtacttg gagcgggtctc tccctccctc atcagcccac caaaccaaac cttagcctcca 4260
agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag aaaatgcctc 4320
caacatgtga ggaagtaatg agagaaatca tagaatttta aggccatgat ttaaggccat 4380
catggcctta atcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 4440
ggcgagcggg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 4500
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 4560
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatacaaaa aatcgacgct 4620
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 4680
gtccctcgt gcgctctcct gttccgacct tgccgcttac cggatacctg tccgccttct 4740
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 4800
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 4860
ccttatccgg taactatcgt cttgagttca acccggttag acacgactta tcgccactgg 4920
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcgggtgt acagagttct 4980
tgaagtgggt gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 5040
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 5100
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 5160
aagaagatcc tttgatcttt tctacggggg ctgacgctca gtggaacgaa aactcacggt 5220
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 5280
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 5340
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 5400
gactcggggg gggggggcgc tgaggtctgc ctctgaaga aggtgttgct gactcatacc 5460
aggcctgaat cgcgccatca tccagccaga aagtgaggga gccacgggtg atgagagctt 5520
tggtgtaggt ggaccagttg gtgattttga acttttgctt tgccacggaa cggctctcgt 5580
tgtcgggaag atgcgtgatc tgatccttca actcagcaaa agttcgattt attcaacaaa 5640
gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtg tacaaccaat taaccaattc 5700
tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatac caggattatc 5760
aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac cgaggcagtt 5820
ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtccaa catcaatata 5880
acctattaat tccccctcgt caaaaataag gttatcaagt gagaaatcac catgagtgc 5940
gactgaatcc ggtgagaatg gcaaaagctt atgcatttct ttccagactt gttcaacagg 6000
ccagccatta cgctcgatc caaaatcact cgcataacc aaaccgttat tcattcgtga 6060
ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa ggacaattac aaacaggaat 6120
cgaatgcaac cggcgcagga aactgccag cgcataaca atattttcac ctgaatcagg 6180
atattcttct aatacctgga atgctgtttt cccggggatc gcagtgggtga gtaaccatgc 6240
atcatcagga gtacggataa aatgcttgat ggtcgggaaga ggcataaatt ccgtcagcca 6300
gtttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc catgtttcag 6360
aaacaactct ggcgcacatcg gcttcccata caatcgatag attgtcgcac ctgattgccc 6420
gacattatcg cgagcccat tatacccata taaatcagca tccatggttg aatttaatcg 6480
cggcctcgag caagacgttt cccgttgaat atggctcata acacccttg tattactggt 6540
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttggt caatgtaaca 6600
tcagagattt tgagacacaa cgtggctttc ccccccccc cattattgaa gcatttatca 6660
gggttattgt ctcagagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 6720
ggttccgcgc acatttcccc gaaaagtgc acctgacgtc taagaaacca ttattatcat 6780
gacattaacc tataaaaaa ggcgatatcac gaggcctttt cgctc 6824

```

&lt;210&gt; 113

&lt;211&gt; 6695

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes



&lt;400&gt; 113

```

tcgcgcgttt  cggatgatgac  ggtgaaaacc  tctgacacat  gcagctcccc  gagacggtca  60
cagcttgtct  gtaagcggat  gccgggagca  gacaagcccc  tcagggcgcg  tcagcgggtg  120
ttggcggtg  tcggggctgg  cttaactatg  cggcatcaga  gcagattgta  ctgagagtgc  180
accatatgcg  gtgtgaaata  ccgcacagat  gcgtaaggag  aaaataccgc  atcagattgg  240
ctattggcca  ttgcatacgt  tgtatccata  tcataatatg  tacatttata  ttggctcatg  300
tccaacatta  ccgccatgtt  gacattgatt  attgactagt  tattaatagt  aatcaattac  360
ggggtcatta  gttcatagcc  catatatgga  gttccgcgtt  acataactta  cggtaaattg  420
ccgcctggc  tgaccgccca  acgacccccg  cccattgacg  tcaataatga  cgtatgttcc  480
catagtaacg  ccaataggga  ctttccattg  acgtcaatgg  gtggagtatt  tacggtaaac  540
tgcccacttg  gcagtacatc  aagtgtatca  tatgccaaag  acgcccccta  ttgacgtcaa  600
tgacggtaaa  tggcccgcc  ggcattatgc  ccagtacatg  accttatggg  actttcctac  660
ttggcagtac  atctacgtat  tagtcacgc  tattaccatg  gtgatgcggt  tttggcagta  720
catcaatggg  cgtggatagc  ggtttgactc  acggggattt  ccaagtctcc  accccattga  780
cgtcaatggg  agtttgttt  ggcacaaaa  tcaacgggac  tttccaaaat  gtcgtaacaa  840
ctccgcccc  ttgacgaaa  tggcggttag  gcgtgtacgg  tgggaggtct  atataagcag  900
agctcgttta  gtgaaccgtc  agatcgcccg  gagacgccat  ccacgtgtt  ttgacctcca  960
tagaagacac  cgggaccgat  ccagcctccg  cggccgggaa  cgggtgcattg  gaacgcggat  1020
tccccgtgcc  aagagtgacg  taagtaccgc  ctatagactc  tataggcaca  cccctttggc  1080
tcttatgcat  gctatactgt  ttttggcttg  gggcctatac  acccccgtt  ccttatgcta  1140
taggtgatgg  tatagcttag  cctatagggt  tgggttattg  accattattg  accactcccc  1200
tattggtgac  gatactttcc  attactaatc  cataacatgg  ctctttgcca  caactatctc  1260
tattggctat  atgccaatat  tctgtccttc  agagactgac  acggactctg  tatttttaca  1320
ggatggggtc  ccatttatta  tttacaaatt  cacatataca  acaacgcgt  cccccgtgcc  1380
cgcagttttt  attaaacata  gcgtgggac  tccacgcgaa  tctcgggtac  gtgttccgga  1440
catgggctct  tctccggtag  cggcgagct  tccacatccg  agccctggtc  ccattgcctc  1500
agcggctcat  ggtcgctcgg  cagctccttg  ctcttaacag  tggaggccag  acttaggcac  1560
agcacaatgc  ccaccaccac  cagtgtgccg  cacaaggccg  tggcggtagg  gtatgtgtct  1620
gaaaatgagc  gtggagattg  ggctcgacg  gctgacgcag  atggaagact  taaggcagcg  1680
gcagaagaag  atgcaggcag  ctgagttgtt  gtattctgat  aagagtcaga  ggtaactccc  1740
gttgcggtgc  tgttaacggt  ggagggcagt  gtagtctgag  cagtactcgt  tgctgccgcg  1800
cgcgccacca  gacataatag  ctgacagact  aacagactgt  tcctttccat  gggctctttc  1860
tgcatgcacc  gtcgtcgaca  cgtgtgatca  gatctcgcg  ccgctctaga  gatctcgcca  1920
ccatgcgcgt  ggagggcac  cagcgcaact  ggaagcagtg  gtggatctgg  ggcacccctg  1980
gcttctggat  ggtgatgatc  tacaacgtgc  gcggcaacct  gtgggtgacc  gtgtactacg  2040
gcgtgcccgt  gtggaaggag  gccaaagacca  ccctgttctg  cgccagcgac  gccaaaggcct  2100
acgacgccga  ggtgcacaac  gtgtgggcca  cccacgcctg  cgtgcccacc  gaccccaacc  2160
cccaggagat  ggtgctggag  aacgtgaccg  agaacttcaa  catgtgggag  aacgacatgg  2220
tgagagcagat  gcaccaggac  atcatcagcc  tgtgggacca  gagcctgaag  ccctgcgtga  2280
agctgacccc  cctgtgcgtg  accctgcact  gcagcaaccg  atcgatcaac  tgcaacacca  2340
gcgccatcac  ccaggcctgc  cccaagggtga  gcttcgaccc  catccccatc  cactactgcg  2400
ccccgcggg  ctacgccatc  ctgaagtgca  acaacaagac  cttcaacggc  accggccccct  2460
gcaacaacgt  gagcaccatc  cagtgcaccc  acggcaccaa  gcccggtgtg  agcaccacg  2520
tgctgctgaa  cggcagcctg  gccgaggagg  agatcatcat  ccgcagcaag  aacctgaccg  2580
acaacgtgaa  gaccatcatc  gtgcacctga  acgagagcgt  ggagatcaac  tgcaccgcgc  2640
ccaacaacaa  caccgcgaag  agcatccgca  tcggccccgg  ccaggccttc  tacgccaccg  2700
gcgagatcat  cggcgacatc  cgccaggccc  actgcaacat  cagccgcacc  gcctggaaca  2760
agaccctgca  ggaggtgggc  aagaagctgg  ccgagcactt  ccccaacaag  gccatcaagt  2820
tcgccaagca  cagcggcggc  gacctggaga  tcaccacca  cagcttcaac  tgccgcggcg  2880
agttcttcta  ctgcaacacc  agcagcctgt  tcaacagcac  ctacaccccc  aacagcaccg  2940
agaacatcac  cggcaccgag  aacagcatca  tcaccatccc  ctgcccgcac  aagcagatca  3000
tcaacatgtg  gcagggcgtg  ggccgcgcca  tgtacgcccc  ccccatcgag  ggcacccctga  3060
cctgcgcgag  caacatcac  ggccgtgtgc  tgaccgcgca  cggcggcacc  ggcacgcacg  3120
acaccgagat  ctccgcgcc  gagggcgggc  acatgcgcga  caactggcgc  agcagactgt  3180

```

```

acaagtacaa ggtggtggag atcaagcccc tgggcatcgc cccacccaag gccaaagctta 3240
ccgtgcaggt gcgccagctg ctgagcggca tcgtgcagca gcagagcaac ctgctgcgcg 3300
ccatcgaggg ccagcagcac atgctgcagc tgaccgtgtg gggcatcaag cagctgcaga 3360
cccgcgtgct ggccatcgag cgctacctgc gcgaccagca gctcgaggac atctggaaca 3420
acatgacctg gatgcagtgg gaccgcgaga tcagcaacta caccaacacc atctaccgcc 3480
tgctggagga cagccagaac cagcaggaga agaacgagca ggacctgctg gccctggaca 3540
agtggcagaa cctgtggacc tggttcggca tcaccaactg gctgtggtac atcaagatct 3600
tcatcaagat cgtggggcggc ctgategggc tgcccatcat cttcgccgtg ctgagcatct 3660
gaacacgtgg gatccagatc tgctgtgcct tctagtgtcc agccatctgt tgtttgcccc 3720
tcccccgctg ctcccttgac cctggaagggt gccactccca ctgtcctttc ctaataaaat 3780
gaggaaattg catcgcatthg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 3840
caggacagca agggggagga ttgggaagac aatagcaggg atgctgggga tgcggtgggc 3900
tctatgggta cccaggtgct gaagaattga cccggttccct cctggggccag aaagaagcag 3960
gcacatcccc ttctctgtga cacacctgt ccacgcccc gggtcttagt tccagcccca 4020
ctcataggac actcatagct caggagggt cgccttcaa tcccaccgc taaagtactt 4080
ggagcggctt ctccctccct catcagccca ccaaaccaaa cctagcctcc aagagtggga 4140
agaaattaaa gcaagatagg ctattaagtg cagagggaga gaaaatgcct ccaacatgtg 4200
aggaagtaat gagagaaatc atagaatttt aaggccatga tttaaggcca tcatggcctt 4260
aatcttccgc ttctctgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 4320
tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa 4380
agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 4440
cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga 4500
ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 4560
tgcgctctcc tgttccgacc ctgccgctta ccgatacct gtccgccttt ctcccttcgg 4620
gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgctt 4680
gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 4740
gtaactatcg tcttgagtcc aaccgggtaa gacacgactt atcgccactg gcagcagcca 4800
ctggtaacag gattagcaga gcgaggtatg taggcgggtg tacagagttc ttgaagtggg 4860
ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag 4920
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 4980
gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 5040
ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 5100
tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 5160
ttaaatcaat ctaaagtata tatgagtaaa cttgggtctga cagttaccaaa tgcttaatca 5220
gtgaggcacc tatctcagcg atctgtctat ttctgttcac catagttgcc tgactcgggg 5280
ggggggggcg ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa 5340
tcgccccatc atccagccag aaagtgaggg agccacgggt gatgagagct ttgttgtagg 5400
tggaaccagt ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa 5460
gatgcgtgat ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgccgtc 5520
ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa 5580
aaactcatcg agcatcaaat gaaactgcaa ttatttcata tcaggattat caataccata 5640
tttttgaaaa agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat 5700
ggcaagatcc tgggtatcgg ctgcgattcc gactcgtcca acatcaatac aacctattaa 5760
tttccccctg tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc 5820
cggtgagaat ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt 5880
acgctcgtca tcaaaatcac tcgcatcaac caaacggtta ttcattcgtg attgcgcctg 5940
agcgagacga aatacgcgat cgctgtttaa aggacaatta caaacaggaa tcgaatgcaa 6000
ccggcgaggg aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc 6060
taatacctgg aatgctgttt tcccggggat cgcagtgggt agtaaccatg catcatcagg 6120
agtacggata aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct 6180
gaccatctca tctgtaacat cattggcaac gctaccttg ccatgtttca gaaacaactc 6240
tggcgcatcg ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc 6300
gcgagcccat ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga 6360
gcaagacgtt tcccgttgaa tatgggtcat aacaccctt gtattactgt ttatgtaagc 6420
agacagtttt attgttcatt atgatataat tttatcttgt gcaatgtaac atcagagatt 6480

```

```

ttgagacaca acgtggcttt ccccccccc ccattattga agcatttatc agggttattg 6540
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 6600
cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac 6660
ctataaaaat aggcgtatca cgaggccctt tcgtc 6695

```

&lt;210&gt; 114

&lt;211&gt; 6614

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 114

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaate cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccggtg tgttaacggg ggaggggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcata cagcgcaact ggaagcagt gtggatctgg ggcacctctg 1980
gcttctggat ggtgatgata tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cgcagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340

```



gcgccatcac	ccaggcctgc	cccaaggtga	gcttcgaccc	catccccatc	cactactgcg	2400
ccccgcggg	ctacgccatc	ctgaagtgc	acaacaagac	cttcaacggc	accggccccct	2460
gcaacaacgt	gagcaccatc	cagtgcaccc	acggcaccaa	gcccgtgggtg	agcaccacagc	2520
tgctgctgaa	cggcagcctg	gccgaggagg	agatcatcat	ccgcagcaag	aacctgaccg	2580
acaacgtgaa	gaccatcatc	gtgcacctga	acgagagcgt	ggagatcaac	tgcacccgcc	2640
gacgggccc	ctgcaacatc	agccgcaccg	cctggaacaa	gaccctgcag	gaggtgggca	2700
agaagctggc	cgagcacttc	cccaacaagg	ccatcaagtt	cgccaagcac	agcggcgggc	2760
acctggagat	caccacccac	agcttcaact	gccgcggcga	gttcttctac	tgcaacacca	2820
gcagcctggt	caacagcacc	tacaccccc	acagcaccga	gaacatcacc	ggcaccgaga	2880
acagcatcat	caccatcccc	tgccgcatac	agcagatcat	caacatgtgg	cagggcggtg	2940
gccgcgccat	gtacgcccc	cccatcgagg	gcatacctgac	ctgccgcagc	aacatcaccg	3000
gcctgctgct	gacccgcgac	ggcggcaccg	gcatacgcga	caccgagatc	ttccgccccg	3060
agggcgggcga	catgcgcgac	aactggcgca	gcgagctgta	caagtacaag	gtggtggaga	3120
tcaagccccct	gggcacgcgc	cccaccaagg	ccaagcttac	cgtgcagggtg	cgccagctgc	3180
tgagcgccat	cgtgcagcag	cagagcaacc	tgctgcgcgc	catcgaggcc	cagcagcaca	3240
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcagac	ccgcgtgctg	gccatcgagc	3300
gctacctgcg	cgaccagcag	ctcgaggaca	tctggaacaa	catgacctgg	atgcagtggg	3360
accgcgagat	cagcaactac	accaacacca	tctaccgcct	gctggaggac	agccagaacc	3420
agcaggagaa	gaacgagcag	gacctgctgg	ccctggacaa	gtggcagaac	ctgtggacct	3480
ggttcggcat	caccaactgg	ctgtggtaca	tcaagatctt	catcaagatc	gtgggcggcc	3540
tgatcggcct	gcgcatacat	ttcgccgtgc	tgagcatctg	aacacgtggg	atccagatct	3600
gctgtgcctt	ctagttgcca	gccatctggt	gtttgccctt	ccccgtgcc	ttccttgacc	3660
ctggaagggtg	ccactccac	tgctctttcc	taataaaatg	aggaaattgc	atcgcatgtg	3720
ctgagtaggt	gtcatctctat	tctggggggt	gggggtggggc	aggacagcaa	gggggaggat	3780
tggaagaca	atagcaggca	tgctggggat	gcggtgggct	ctatgggtac	ccaggtgctg	3840
aagaattgac	ccggttcctc	ctggggccaga	aagaagcagg	cacatccccct	tctctgtgac	3900
acaccctgtc	cacgccccctg	gttcttagtt	ccagccccac	tcataggaca	ctcatagctc	3960
aggagggctc	cgccttcaat	cccacccgct	aaagtacttg	gagcgggtctc	ttcctccctc	4020
atcagcccac	caaaccacac	ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	4080
tattaagtgc	agaggagag	aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	4140
tagaatttta	aggccatgat	ttaaggccat	catggcctta	atcttccgct	ttctcgctca	4200
ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	tcaaaggcgg	4260
taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	4320
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	4380
cccctgacga	gcatacaaaa	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	4440
tataaagata	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	4500
tgccgcttac	cggatacctg	tcgccttttc	ttccttcggg	aagcgtggcg	ctttctcata	4560
gctcacgctg	taggtatctc	agttcggtgt	aggctcgttcg	ctccaagctg	ggctgtgtgc	4620
acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	4680
accgggtaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	4740
cgagggtatgt	aggcggtgct	acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	4800
gaagaacagt	atcttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	4860
gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	4920
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	4980
ctgacgctca	gtggaacgaa	aactcacggt	aagggaatttt	ggtcatgaga	ttatcaaaaa	5040
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	5100
atgagtaaac	ttgggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	5160
tctgtctatt	tcgttcaccc	atagttgcct	gactcggggg	ggggggggcg	tgaggtctgc	5220
ctcgtgaaga	aggtgttgct	gactcatacc	aggcctgaat	cgccccatca	tccagccaga	5280
aagtgaggga	gccacgggtg	atgagagctt	tggtgtagggt	ggaccagttg	gtgattttga	5340
acttttgctt	tgccacggaa	cggctctgct	tgctcggaag	atgcgtgac	tgatccttca	5400
actcagcaaa	agttcgattt	attcaacaaa	gccgccgtcc	cgtcaagtca	gcgtaatgct	5460
ctgccagtgt	tacaaccaat	taaccaattc	tgattagaaa	aactcatcga	gcatacaaat	5520
aaactgcaat	ttattcatat	caggattatc	aataccatat	ttttgaaaaa	gccgtttctg	5580
taatgaagga	gaaaactcac	cgaggcagtt	ccataggatg	gcaagatcct	ggtatcggtc	5640

```

tgcgattccg actcgtccaa catcaataca acctattaat ttccccctcgt caaaaataag 5700
gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt 5760
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact 5820
cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgac 5880
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 5940
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 6000
cccggggac gcagtgggtga gtaacctatgc atcatcagga gtacggataa aatgcttgat 6060
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 6120
attggcaacg ctacctttgc catgtttcag aaacaactct ggccgcatcgg gcttcccata 6180
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 6240
taaatacagca tccatgttgg aatttaatcg cggcctcgag caagacgttt ccggttgaat 6300
atggctcata acacccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 6360
tgatataatt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 6420
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 6480
tgaatgtatt tagaaaaata aacaaatagg gggtccgcgc acatttcccc gaaaagtgcc 6540
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6600
gaggcccttt cgtc 6614

```

&lt;210&gt; 115

&lt;211&gt; 6545

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 115

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggctc atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgaggttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgcctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560

```



```

agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat ggtcttttcc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcggt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacccctg 1980
gcttctggat ggtgatgac tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccggt gtggaaggag gccaaagacca ccctgttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgatcaac tgcaaacacca 2340
gcgccatcac ccaggcctgc cccaagggtga gcttcgaccc catccccatc cactactgcg 2400
cccccgccgg ctacgccatc ctgaagtgc acaacaagac cttcaacggc accggccccct 2460
gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccggtggtg agcaccagc 2520
tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aacctgaccg 2580
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcacccgcc 2640
gatcgggcca ctgcaacatc agccgcaccg cctggaacaa gaccctgcag gaggtgggca 2700
agaagctggc cgagcacttc cccaacaagg ccatcaagtt cgccaagcac agcggcggcg 2760
acctggagat caccacccac agcttcaact gccgcggcga gttcttctac tgcaaacacc 2820
gatcgatccc ctgccgcac aagcagatca tcaacatgtg gcaggcgctg ggccgcgcca 2880
tgtacgcccc ccccatcgag ggcacccctga cctgccgcag caacatcacc ggccgtctgc 2940
tgacccgcga cggcggcacc ggcacgcag acacgcagat cttccgcccc gagggcgggcg 3000
acatgcgcga caactggcgc agcgagctgt acaagtacaa ggtggtggag atcaagcccc 3060
tgggcatcgc cccaccaag gccaaagctta ccgtgcaggt gcgccagctg ctgagcggca 3120
tcgtgcagca gcagagcaac ctgctgcgcg ccatecgaggc ccagcagcac atgctgcagc 3180
tgaccgtgtg gggcatcaag cagctgcaga cccgcgtgct ggccatcgag cgctacctgc 3240
gcgaccagca gctcgaggac atctggaaca acatgacctg gatgcagtgg gaccgcgaga 3300
tcagcaacta caccaacacc atctaccgcc tgctggagga cagccagaac cagcaggaga 3360
agaacgagca ggacctgctg gccctggaca agtggcagaa cctgtggacc tgggtcggca 3420
tcaccaactg gctgtggtac atcaagatct tcatcaagat cgtgggcggc ctgatcggcc 3480
tgcgcatcat ctgcgcgtg ctgagcatct gaacacgtgg gatccagatc tgctgtgect 3540
tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaagggt 3600
gccactccca ctgtccttcc ctaataaaaat gaggaaattg catcgcatct tctgagtagg 3660
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggaggga ttgggaagac 3720
aatagcaggc atgctgggga tgcggtgggc tctatgggta cccagggtgct gaagaattga 3780
cccggttcct cctgggccag aaagaagcag gcacatcccc ttctctgtga cacaccctgt 3840
ccacgcccct ggttcttagt tccagcccca ctcataggac actcatagct caggagggt 3900
ccgccttcaa tcccaccgc taaagtactt ggagcggctt ctcctccct catcagccca 3960
ccaaacaaaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagt 4020
cagagggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4080
aaggccatga ttttaaggcca tcatggcctt aatcttccgc ttctctgctc actgactcgc 4140
tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg 4200
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 4260
ccaggaaccg taaaaaggcc gcgttgcctg cgtttttcca taggctccgc cccctgacg 4320
agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4380
accaggcgtt tccccctgga agctccctcg tgcgctctcc tgbtccgacc ctgccgctta 4440
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4500
gtaggtatct cagttcggtg taggtcgctt gctccaagct gggctgtgtg cacgaacccc 4560
ccgttcagcc cgaccgctgc gccttatecg gtaactatcg tcttgagtcc aaccggtaa 4620
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4680
taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag 4740
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 4800
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 4860

```

```

cgcgagagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 4920
agtggaaacga aaactcacgt taagggatct tggatcatgag attatcaaaa aggatcttca 4980
cctagatcct tttaaattaa aaatgaagtt ttaaatacaat cttaaagtata tatgagtaaa 5040
cttgggtctga cagttaccaaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat 5100
ttcgtttcate catagttgcc tgactcgggg gggggggggcg ctgaggtctg cctcgtgaag 5160
aaggtgttgc tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaagg 5220
agccacgggt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5280
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgaccttc aactcagcaa 5340
aagttcgatt tattcaacaa agccgcctgc ccgtcaagtc agcgtaatgc tctgccagt 5400
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaa at gaaactgcaa 5460
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5520
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggatcgggt ctgcgattcc 5580
gactcgtcca acatcaatac aacctattaa tttccccctc tcaaaaataa ggttatcaag 5640
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaaagct tatgcatttc 5700
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 5760
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 5820
aggacaatta caaacaggaa tcgaatgcaa ccggcgagc aacactgcca gcgcatcaac 5880
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 5940
cgagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6000
aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6060
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6120
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6180
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6240
aacaccctt gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt 6300
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6360
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6420
ttagaaaaat aaacaaatag ggggtccgcg cacatttccc cgaaaagtgc cacctgacgt 6480
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 6540
tcgtc 6545

```

&lt;210&gt; 116

&lt;211&gt; 6626

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 116

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgcattcata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctgg tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840

```

```

ctccgcccc ttgacgcaaa tgggcggttag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaate cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatgggggc ccattttatta tttaaaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggctcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtctcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcgaact ggaagcagt gtggatctgg ggcatectgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaagacca ccctgttctg cgcagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgatcaac tgcaacacca 2340
gcgccatcac ccaggcctgc cccaagggtga gcttcgaccc catccccatc cactactgcg 2400
ccccgcgg ctacgccatc ctgaagtga acaacaagac cttaacggc accggcccct 2460
gcaacaacgt gagcaccatc cagtgcaccc acggcaccaa gcccggtggg agcaccacgc 2520
tgctgctgaa cggcagcctg gccgaggagg agatcatcat ccgcagcaag aactgaccg 2580
acaacgtgaa gaccatcatc gtgcacctga acgagagcgt ggagatcaac tgcaccgcgc 2640
ccaacaaca caccgcaag agcatccgca tcggccccgg ccaggccttc tacgccaccg 2700
gcgagatcat cggcgacatc cgccaggccc actgcaacat cagccgcacc gcctggaaca 2760
agaccctgca ggaggtgggc aagaagctgg ccgagcactt cccaacaag gccatcaagt 2820
tcgccaagca cagcggcggc gacctggaga tcaccacca cagcttcaac tgccgcggcg 2880
agttcttcta ctgcaacacc cgatcgatcc cctgcgcgat caagcagatc atcaacatgt 2940
ggcagggcgt gggccgcgcc atgtacgccc ccccatcga gggcatcctg acctgccgca 3000
gcaacatcac cggcctgctg ctgaccgcgc acggcggcac cggcatgcac gacaccgaga 3060
tcttcgccc cgagggcggc gacatgcgcg acaactggcg cagcgagctg tacaagtaca 3120
aggtggtgga gatcaagccc ctgggcatcg ccccaacca ggccaagctt accgtgcagg 3180
tgccccagct gctgagcggc atcgtgcagc agcagagcaa cctgctgcgc gccatcgagg 3240
cccagcagca catgctgcag ctgaccgtgt ggggcatcaa gcagctgcag acccgcgctg 3300
tggccatcga gcgctacctg cgcgaccagc agctcgagga catctggaac aacatgacct 3360
ggatgcagtg ggaccgcgag atcagcaact acaccaacac catctaccgc ctgctggagg 3420
acagccagaa ccagcaggag aagaacgagc aggacctgct ggccctggac aagtggcaga 3480
acctgtggac ctggttcggc atcaccaact ggctgtggta catcaagatc ttcacataga 3540
tcgtgggcgg cctgatcggc ctgcgcatca tcttcgcccgt gctgagcatc tgaacacgtg 3600
ggatccagat ctgctgtgcc ttctagtgtc cagccatctg ttgtttgccc ctccccctg 3660
ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt 3720
gcatcgcat gtctgagtag gtgtcattct attctggggg gtgggggtggg gcaggacagc 3780
aagggggagg attgggaaga caatagcagg catgctgggg atgcgggtggg ctctatgggt 3840
accaggtgc tgaagaattg acccggttcc tcctgggcca gaaagaagca ggcacatccc 3900
cttctctgtg acacaccctg tccacgcccc tgggtcttag ttccagcccc actcatagga 3960
cactcatagc tcaggagggc tccgccttca atcccacccg ctaaagtact tggagcggtc 4020
tctccctccc tcatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa 4080
agcaagatag gctattaagt gcagaggggag agaaaatgcc tccaacatgt gaggaagtaa 4140

```



```

tgagagaaat catagaatth taaggccatg atttaaggcc atcatggcct taatcttccg 4200
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct ggcgcgagcg gtatcagctc 4260
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 4320
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 4380
ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 4440
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc 4500
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 4560
cgctttctca tagctcacgc tgtaggatc tcagttcggg gtaggctcgt cgctccaagc 4620
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 4680
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca 4740
ggattagcag agcgaggat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact 4800
acggctacac tagaagaaca gtatttggtt tctgcgctct gctgaagcca gttaccttcg 4860
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt 4920
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 4980
tttctacggg gtctgacgct cagtggaaac aaaactcacg ttaagggttt ttggtcatga 5040
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaataca 5100
tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaatc agtgaggcac 5160
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcggg gggggggggc 5220
gctgaggtct gcctcgtgaa gaagggtgtt ctgactcata ccaggcctga atcgccccat 5280
catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt 5340
tgggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga 5400
tctgatcctt caactcagca aaagtctgat ttattcaaca aagccgccgt cccgtcaagt 5460
cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcatc 5520
gagcatcaaa tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa 5580
aagccgtttc tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc 5640
ctgggtatcgg tctgcgattc cgactcgtcc aacatcaata caacctatta atttccctc 5700
gtcaaaaata aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa 5760
tggcaaaagc ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc 5820
atcaaaatca ctgcacatca ccaaaccgtt attcattcgt gattgcgcct gagcgagacg 5880
aaatacgcga tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgagc 5940
gaacactgcc agcgcatcaa caatatthtc acctgaatca ggatattctt ctaataacctg 6000
gaatgctgtt ttcccgggga tcgcagtggg gagtaaccat gcacatcag gagtacggat 6060
aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc 6120
atctgtaaca tcattggcaa cgctacctt gccatgttcc agaaacaact ctggcgcatc 6180
gggcttccca tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca 6240
tttatacca tataaatcag catccatgtt ggaatttaac cgcggcctcg agcaagacgt 6300
ttcccgttga atatggctca taacacccct tgtattactg tttatgtaag cagacagttt 6360
tattgttcat gatgatatat ttttatcttg tgcaatgtaa catcagagat tttgagacac 6420
aacgtggctt tccccccccc cccattattg aagcatttat cagggttatt gtctcatgag 6480
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc 6540
ccgaaaagtg ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa 6600
taggogtate acgaggccct ttcgtc 6626

```

&lt;210&gt; 117

&lt;211&gt; 6743

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 117

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60

```

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggcttttcc 1860
tgagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcac cagcgcgaact ggaagcagtg gtggatctgg ggcacccctg 1980
gcttctggat ggtgatgac tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cgcagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgaagaac tgcagcttca 2340
acatgaccac cgaggtgcgc gacaagcgcg agaaggtgca cgcctgttcc taccgcctgg 2400
acatcgtgcc cctgaagaac gagagcagca acaccagcgg cgactaccgc ctgatcaact 2460
gcaacaccag cgccatcacc caggcctgcc ccaaggtgag cttcgacccc atccccatcc 2520
actactgcgc ccccgccggc tacgccatcc tgaagtgcaa caacaagacc ttcaacggca 2580
ccggccctg caacaacgtg agcaccatcc agtgaccca cggcaccaag cccgtgggtga 2640
gcaccagct gctgctgaac ggcagcctgg ccgaggagga gatcatcatc cgcagcaaga 2700
acctgaccga caacgtgaag accatcatcg tgcacctgaa cgagagcgtg gagatcaact 2760
gcaccgcgg atcggccccc tgcaacatca gccgcaccgc ctggaacaag accctgcagg 2820
aggtgggcaa gaagctggcc gagcacttcc ccaacaaggc catcaagtcc gccaaagcaca 2880
gcggcgggca cctggagatc accaccaca gcttcaactg ccgcggcgag ttcttctact 2940
gcaacaccag cagcctgttc aacagcacct acaccccaaa cagcaccgag aacatcaccg 3000
gcaccgagaa cagcatcatc accatccct gccgcacaa gcagatcatc aacatgtggc 3060
agggcgtgg ccgcgccatg tacgcccccc ccatcgaggg catcctgacc tgccgcagca 3120
acatcaccgg cctgtgctg acccgcgacg gcggcaccgg catgcacgac accgagatct 3180
tccgccccga gggcgggcag atgcgcgaca actggcgag cgagctgtac aagtacaagg 3240
tgggtggagat caagcccctg ggcacgccc ccaccaaggc caagcttacc gtgcagggtc 3300
gccagctgct gagcggcatc gtgcagcagc agagcaacct gctgcgcgcc atcgaggccc 3360

```



```

agcagcacat gctgcagctg accgtgtggg gcatcaagca gctgcagacc cgcgtgctgg 3420
ccatcgagcg ctacctgcgc gaccagcagc tcgaggacat ctggaacaac atgacctgga 3480
tgcagtggga ccgcgagatc agcaactaca ccaacacat ctaccgctg ctggaggaca 3540
gccagaacca gcaggagaag aacgagcagg acctgctggc cctggacaag tggcagaacc 3600
tgttgacctg gttcggcatc accaactggc tgttggtacat caagatcttc atcaagatcg 3660
tgggcggcct gatcggcctg cgcattcatc tcgccgtgct gagcatctga acacgtggga 3720
tccagatctg ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgctgct 3780
tccttgacct tgggaagggtgc cactcccact gtccttttct aataaaatga ggaaattgca 3840
tcgcattgtc tgagtaggtg tcatcttatt ctgggggggtg ggggtggggca ggacagcaag 3900
ggggaggatt gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgggtacc 3960
caggtgctga agaattgacc cggttcctcc tgggccagaa agaagcaggc acatcccctt 4020
ctctgtgaca caccctgtcc acgcccctgg ttcttagttc cagcccact cataggacac 4080
tcatagctca ggagggtctc gccttcaatc ccaccgcta aagtacttgg agcgtctct 4140
ccctccctca tcagcccacc aaaccaaacc tagcctccaa gagtgggaag aaattaaagc 4200
aagataggct attaatgtga gaggagaga aaatgcctcc aacatgtgag gaagtaatga 4260
gagaaatcat agaattttaa ggccatgatt taaggccatc atggccttaa tcttcgctt 4320
cctcgctcac tgactcgctg cgctcggctg ttcggctgct gcgagcggta tcagctcact 4380
caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag 4440
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 4500
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 4560
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgct cgctctcctg 4620
ttccgacctc gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcg 4680
tttctcatag ctacgctgt aggtatctca gttcgggtga ggtcggtcgc tccaagctgg 4740
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgct 4800
ttgagtccaa ccggttaaga cagcacttat cgccactggc agcagccact ggtaacagga 4860
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg 4920
gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 4980
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg 5040
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 5100
ctacgggggtc tgacgctcag tggaaacgaaa actcacgtta agggattttg gtcattgagat 5160
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 5220
aaagtatata tgagtaaact tggctctgaca gttaccaatg cttaatcagt gaggcacct 5280
tctcagcgat ctgtctattt cgttcaccca tagttgcctg actcgggggg ggggggcgct 5340
gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc gcccacatcat 5400
ccagccagaa agtgaggag ccacgggtga tgagagcttt gttgtagggt gaccagttgg 5460
tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga tgcgtgatct 5520
gatccttcaa ctacgcaaaa gttcgattta ttcaacaaag ccgcgctccc gtcaagtcag 5580
cgtaatgctc tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag 5640
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag 5700
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg 5760
gtatcgggtc gcgattccga ctcgctccaac atcaatacaa cctattaatt tcccctcgct 5820
aaaaataagg ttatcaagt agaaatcacc atgagtgcag actgaatccg gtgagaatgg 5880
caaaagctta tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcac 5940
aaaatcactc gcatcaacca aaccgttatt cattcgatg tgcgcctgag cgagacgaaa 6000
tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgaggaa 6060
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa 6120
tgctgttttc ccggggatcg cagtgggtgag taaccatgca tcatcaggag tacggataaa 6180
atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc 6240
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg 6300
cttcccatac aatcgataga ttgtcgcacc tgattgcccc acattatcgc gagccatttt 6360
atacccatat aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc 6420
ccgttgaaata tggctcataa cacccttgtt attactgttt atgtaagcag acagttttat 6480
tgttcatgat gatataatctt tatcttgtgc aatgtaacat cagagatttt gagacacaac 6540
gtggcttttc ccccccccc attattgaag catttatcag ggttattgtc tcatgagcgg 6600
atacatatct gaatgtatct agaaaaataa acaaataggg gttccgcgca catttccccg 6660

```

aaaagtgcc cctgacgtct aagaaacccat tattatcatg acattaacct ataaaaatag 6720  
gcgtatcacg aggccctttc gtc 6743

<210> 118

<211> 6674

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 118

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420  
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480  
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600  
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttccctac 660  
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac ttcccaaaat gtcgtaacaa 840  
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt ttgacctcca 960  
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380  
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
agcggctcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920  
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacccctg 1980  
gcttctggat ggtgatgata tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040  
gcgtgcccgt gtggaaggag gccaaagcca ccctgttctg cgcacgcgac gccaaaggcct 2100  
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160  
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220  
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280  
agctgacccc cctgtgcgtg accctgcact gcagcaaccg atcgaagaac tgcagcttca 2340  
acatgaccac cgaggtgcgc gacaagcgcg agaagggtgca cgccctgttc taccgcctgg 2400  
acatcgtgcc cctgaagaac gagagcagca acaccagcgg cgactaccgc ctgatcaact 2460

gcaacaccag cgccatcacc caggcctgcc ccaaggtgag cttegacccc atccccatcc 2520  
actactgcgc ccccgccggc tacgccatcc tgaagtgcaa caacaagacc ttcaacggca 2580  
ccggccccctg caacaacgtg agcaccatcc agtgcaccca cggcaccaag cccgtggtga 2640  
gcaccagct gctgctgaac ggcagcctgg ccgaggagga gatcatcatc cgcagcaaga 2700  
acctgaccga caacgtgaag accatcatcg tgcacctgaa cgagagcgtg gagatcaact 2760  
gcaccgcgcg atcgccccac tgcaacatca gccgcaccgc ctggaacaag accctgcagg 2820  
aggtgggcaa gaagctggcc gagcacttcc ccaacaaggc catcaagtcc gccaaagcaca 2880  
gcggcggcga cctggagatc accacccaca gcttcaactg ccgcgggcag ttcttctact 2940  
gcaacacccg atcgatcccc tgccgcatca agcagatcat caacatgtgg cagggcgtgg 3000  
gccgcgccat gtacgcccc ccctcgcagg gcatcctgac ctgccgcagc aacatcaccg 3060  
gcctgctgct gaccgcgcac ggcggcaccg gcatgcacga caccgagatc ttccgccccg 3120  
agggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag gtggtggaga 3180  
tcaagcccct gggcatcgcc cccaccaagg ccaagcttac cgtgcagggtg cgccagctgc 3240  
tgagcggcat cgtgcagcag cagagcaacc tgctgcgcgc catcgaggcc cagcagcaca 3300  
tgctgcagct gaccgtgtgg ggcataaagc agctgcagac ccgctgctg gccatcgagc 3360  
gctacctgcg cgaccagcag ctcgaggaca tctggaacaa catgacctgg atgcagtggg 3420  
accgcgagat cagcaactac accaacacca tctaccgctt gctggaggac agccagaacc 3480  
agcaggagaa gaacgagcag gacctgctgg ccctggacaa gtggcagAAC ctgtggacct 3540  
gggtcggcat caccaactgg ctgtggtaca tcaagatctt catcaagatc gtgggcggcc 3600  
tgatcggcct gcgcatcatc ttccgctgc tgagcatctg aacacgtggg atccagatct 3660  
gctgtgcctt ctagtggcca gccatctgtt gtttgccctt ccccgctgcc ttcttgacc 3720  
ctggaagggtg ccactccac tgctcttctt taataaaatg aggaaattgc atcgattgt 3780  
ctgagtaggt gtcattctat tctgggggtt ggggtggggc aggacagcaa gggggaggat 3840  
tggaagaca atagcaggca tgctggggat gcggtgggct ctatgggtac ccagggtctg 3900  
aagaattgac ccggttcctc ctgggccaga aagaagcagg cacatcccct tctctgtgac 3960  
acacctgtc cagccccctg gttcttagtt ccagccccac tcataggaca ctcatagctc 4020  
aggagggtc cgccttcaat cccaccgct aaagtacttg gagcggtctc tccctccctc 4080  
atcagcccac caaaccaaac ctagcctcca agagtgggaa gaaattaaag caagataggc 4140  
tattaagtgc agaggagag aaaatgcctc caacatgtga ggaagtaatg agagaaatca 4200  
tagaatttta aggccatgat ttaaggccat catggcctta atcttccgct tctcgtca 4260  
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg 4320  
taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 4380  
agcaaaaggc caggaaccgt aaaaaggcgg cgttgctggc gtttttccat aggctccgcc 4440  
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 4500  
tataaagata ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacc 4560  
tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata 4620  
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 4680  
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 4740  
acccggtaa agcagactta tcgccactgg cagcagccac tggtaacagg attagcagag 4800  
cgaggtatgt aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 4860  
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 4920  
gtagctcttg atccggcaaa caaacaccg ctggtagcgg tgggtttttt gtttgcaagc 4980  
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 5040  
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 5100  
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 5160  
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 5220  
tctgtctatt tcgttcatcc atagttgcct gactcggggg gggggggcgc tgaggctctgc 5280  
ctcgtgaaga aggtgttgct gactcatacc aggcctgaat cggcccatca tccagccaga 5340  
aagtgaggga gccacgggtg atgagagctt tgttgtaggt ggaccagttg gtgattttga 5400  
acttttgctt tgccacggaa cggctctgct tgcgggaag atgcgtgatc tgatccttca 5460  
actcagcaaa agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct 5520  
ctgccagtg tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg 5580  
aaactgcaat ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg 5640  
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc 5700  
tgcgattccg actcgtccaa catcaataca acctattaat ttcccctcgt caaaaataag 5760



```

gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt 5820
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgatcat caaaatcact 5880
cgcatcaacc aaaccggttat tcatttcgtga ttgcgccctga gcgagacgaa atacgcgac 5940
gctgttataaa ggacaattac aaacaggaat cgaatgcaac cggcgagga acactgccag 6000
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 6060
cccggggatc gcagtgggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 6120
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 6180
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcacggt gcttcccata 6240
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 6300
taaatcagca tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat 6360
atggctcata acacccttg tattactgtt tatgtaagca gacagtttta ttgttcacga 6420
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 6480
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 6540
tgaatgtatt tagaaaaata aacaaatagg ggttcgcgc acatttcccc gaaaagtgcc 6600
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 6660
gaggcccttt cgtc 6674

```

&lt;210&gt; 119

&lt;211&gt; 6755

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 119

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcggtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aagacccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcccg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgaac taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620

```

gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgata	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaaagacca	ccctgttctg	cgccagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	atcgaagaac	tgacgcttca	2340
acatgaccac	cgaggtgcgc	gacaagcgcg	agaaggtgca	cgccctgttc	taccgcctgg	2400
acatcgtgcc	cctgaagaac	gagagcagca	acaccagcgg	cgactaccgc	ctgatcaact	2460
gcaacaccag	cgccatcacc	caggcctgcc	ccaaggtgag	cttcgacccc	atccccatcc	2520
actactgcgc	ccccgccggc	tacgccatcc	tgaagtgcaa	caacaagacc	ttcaacggca	2580
ccggcccctg	caacaacgtg	agcaccatcc	agtgcaccca	cggcaccaag	cccgtggtga	2640
gcacccagct	gctgctgaac	ggcagcctgg	ccgaggagga	gatcatcatc	cgcagcaaga	2700
acctgaccga	caacgtgaag	accatcatcg	tgcacctgaa	cgagagcgtg	gagatcaact	2760
gcacccgccc	caacaacaac	acccgcaaga	gcatccgcct	cggccccggc	caggccttct	2820
acgccaccgg	cgagatcatc	ggcgacatcc	gccaggccca	ctgcaacatc	agccgcaccg	2880
cctggaacaa	gaccctgcag	gaggtgggca	agaagctggc	cgagcacttc	cccaacaagg	2940
ccatcaagtt	cgccaagcac	agcggcgggc	acctggagat	caccacccac	agcttcaact	3000
gccgcggcga	gttctttctac	tgaacacccc	gatcgatccc	ctgccgcctc	aagcagatca	3060
tcaacatgtg	gcagggcgtg	ggccgcgcca	tgtacgcccc	ccccatcgag	ggcatcctga	3120
cctgccgcag	caacatcacc	ggcctgctgc	tgacccgcga	cggcggcacc	ggcatgcacg	3180
acaccgagat	cttccgcccc	gagggcgggc	acatgcgcga	caactggcgc	agcgagctgt	3240
acaagtacaa	ggtggtggag	atcaagcccc	tgggcatcgc	ccccaccaag	gccaagctta	3300
ccgtgcaggt	gcgccagctg	ctgagcggca	tcgtgcagca	gcagagcaac	ctgctgcgcg	3360
ccatcgaggc	ccagcagcac	atgctgcagc	tgaccgtgtg	gggcatcaag	cagctgcaga	3420
cccgcgtgct	ggccatcgag	cgctacctgc	gcgaccagca	gctcgaggac	atctggaaca	3480
acatgacctg	gatgcagtg	gaccgcgaga	tcagcaacta	caccaacacc	atctaccgcc	3540
tgctggagga	cagccagaac	cagcaggaga	agaacgagca	ggacctgctg	gccctggaca	3600
agtggcagaa	cctgtggacc	tgggttcggca	tcaccaactg	gctgtggtac	atcaagatct	3660
tcataagat	cgtgggcggc	ctgatcggcc	tgcgcctcat	cttcgccgtg	ctgagcatct	3720
gaacacgtgg	gatccagatc	tgctgtgect	tctagtgtgc	agccatctgt	tgtttgcccc	3780
tcccccgctg	cttccttgac	cctggaagg	gccactccca	ctgtcccttc	ctaataaaat	3840
gaggaaattg	catcgcattg	tctgagtagg	tgtcattcta	ttctgggggg	tggggtgggg	3900
caggacagca	agggggagga	ttgggaagac	aatagcaggc	atgctgggga	tgcggtgggc	3960
tctatgggta	cccaggtgct	gaagaattga	cccggttcct	cctgggccag	aaagaagcag	4020
gcacatcccc	ttctctgtga	cacaccctgt	ccacgcccct	ggttcttagt	tccagcccca	4080
ctcataggac	actcatagct	caggagggtc	ccgccttcaa	tcccaccgcg	taaagtactt	4140
ggagcgggtc	ctccctccct	catcagccca	ccaaacccaa	cctagcctcc	aagagtggga	4200
agaaattaaa	gcaagatagg	ctattaagtg	cagagggaga	gaaaatgcct	ccaacatgtg	4260
aggaagtaat	gagagaaatc	atagaatttt	aaggccatga	tttaaggcca	tcattggcct	4320
aatcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	4380
tatcagctca	ctcaaaggcg	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	4440
agaacatgtg	agcaaaaagg	cagcaaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	4500
cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtacga	4560
ggtggcgaaa	cccagacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	4620
tgcgctctcc	tggtccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	4680
gaagcgtggc	gctttctcat	agctcacgct	gtaggatatc	cagttcggtg	taggtcggtc	4740
gctccaagct	gggctgtgtg	cacgaacccc	cgttcagccc	cgaccgctgc	gccttatccg	4800
gtaactatcg	tcttgagtc	aaccgggtaa	gacacgactt	atcgccactg	gcagcagcca	4860
ctggtaacag	gattagcaga	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtgg	4920



```

ggcctaacta cggctacact agaagaacag tatattggtat ctgcgctctg ctgaagccag 4980
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 5040
gtgggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 5100
ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt 5160
tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 5220
ttaaatcaat cttaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca 5280
gtgaggcacc tatctcagcg atctgtctat ttcgttcate catagttgcc tgactcgggg 5340
ggggggggcg ctgagggtctg cctcgtgaag aagggtgtgc tgactcatac caggcctgaa 5400
tcgccccatc atccagccag aaagtgaagg agccacggtt gatgagagct ttgtttagg 5460
tggaaccagt ggtgattttg aacttttgc tttgccacga acggtctgcg ttgtcgggaa 5520
gatgcgtgat ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgccgtc 5580
ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa 5640
aaactcatcg agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata 5700
tttttgaaaa agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat 5760
ggcaagatcc tgggtatcgg tgcgattcc gactcgtcca acatcaatac aacctattaa 5820
tttccccctc tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc 5880
cggtgagaat ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt 5940
acgctcgtca tcaaaatcac tcgcatcaac caaacggtta ttcattcgtg attgcgcctg 6000
agcgagacga aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa 6060
ccggcgcgagg aacactgcca gcgcatcaac aatatatttc cctgaatcag gatattcttc 6120
taatacctgg aatgctgttt tcccggggat cgcagtgggt agtaaccatg catcatcagg 6180
agtacggata aatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct 6240
gaccatctca tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc 6300
tggcgcatcg ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc 6360
gcgagcccat ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga 6420
gcaagacgtt tcccgttgaa tatggctcat aacaccctt gtattactgt ttatgtaagc 6480
agacagtttt attgttcatg atgatataatt tttatcttgt gcaatgtaac atcagagatt 6540
ttgagacaca acgtggcttt ccccccccc ccattattga agcatttatc agggttattg 6600
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 6660
cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac 6720
ctataaaaat aggcgtatca cgaggccctt tcgtc 6755

```

&lt;210&gt; 120

&lt;211&gt; 6677

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 120

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc tggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcategc tattaccatg gtgatgcggt tttggcagta 720

```

catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtataca	840
ctccgcccc	ttgacgcaa	tggcggttag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggtagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgcccg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgatc	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaaagcca	ccctgttctg	cgcacgcgac	gccaaggcct	2100
acgacgccga	ggtgcacaa	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	ccgatcgatc	aactgcaaca	2400
ccagcgccat	caccagggcc	tgcaccaagg	tgagcttcga	ccccatcccc	atccactact	2460
gcgccccgc	cggctacgcc	atcctgaagt	gcaacaacaa	gaccttcaac	ggcaccggcc	2520
cctgcaacaa	cgtgagcacc	atccagtgc	cccacggcac	caagcccgtg	gtgagcacc	2580
agctgctgct	gaacggcagc	ctggccgagg	aggagatcat	catccgcagc	aagaacctga	2640
ccgacaacgt	gaagaccatc	atcgtgcacc	tgaacgagag	cgtggagatc	aactgcacc	2700
gccgatcggc	ccactgcaac	atcagccgca	ccgcctggaa	caagaccctg	caggaggtgg	2760
gcaagaagct	ggccgagcac	ttccccaaca	aggccatcaa	gttcgccaag	cacagcggcg	2820
gcgacctgga	gatcaccacc	cacagcttca	actgccgcgg	cgagtctctc	tactgcaaca	2880
ccagcagcct	gttcaacagc	acctacacc	ccaacagcac	cgagaacatc	accggcaccg	2940
agaacagcat	catcaccatc	ccctgccgca	tcaagcagat	catcaacatg	tggcagggcg	3000
tgggcgcgc	catgtacgcc	cccccatcg	agggcattct	gacctgccgc	agcaacatca	3060
ccggcctgct	gctgaccgcg	gacggcggca	ccggcatgca	cgacaccgag	atcttccgcc	3120
ccgagggcgg	cgacatgcgc	gacaactggc	gcagcgagct	gtacaagtac	aaggtggtgg	3180
agatcaagcc	cctgggcate	gccccacca	aggccaagct	taccgtgcag	gtgcgccagc	3240
tgctgagcgg	catcgtgcag	cagcagagca	acctgctgcg	cgccatcgag	gcccagcagc	3300
acatgctgca	gctgaccgtg	tggggcatca	agcagctgca	gacccgcgtg	ctggccatcg	3360
agcgtacct	gcgcgaccag	cagctcgagg	acatctggaa	caacatgacc	tggatgcagt	3420
gggaccgcga	gatcagcaac	tacaccaaca	ccatctaccg	cctgctggag	gacagccaga	3480
accagcagga	gaagaacgag	caggacctgc	tggccctgga	caagtggcag	aacctgtgga	3540
cctggttcgg	catcaccaac	tggctgtggt	acatcaagat	cttcatcaag	atcgtggggc	3600
gcctgatecg	cctgcgcate	atcttcgccc	tgtgagcat	ctgaacacgt	gggatccaga	3660
tctgctgtgc	cttctagtgt	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	3720
acctggaag	gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcacgcgat	3780
tgtctgagta	ggtgtcatte	tattctgggg	ggtgggggtg	ggcaggacag	caagggggag	3840
gattgggaag	acaatagcag	gcattgctgg	gatgcgggtg	gctctatggg	taccaggtg	3900
ctgaagaatt	gacccggttc	ctcctggggc	agaaagaagc	aggcacatcc	ccttctctgt	3960
gacacaccct	gtccacgccc	ctggttctta	gttccagccc	cactcatagg	acactcatag	4020

```

ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggg ctctccctcc 4080
ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4140
ggctattaag tgcagagggg gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4200
tcatagaatt ttaaggccat gatttaagge catcatggcc ttaatcttcc gcttcctcgc 4260
tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4320
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 4380
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4440
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4500
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 4560
ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 4620
atagctcacg ctgtaggtat ctgagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 4680
tgcacgaacc cccggttcag cccgaccgct gcgccttate cggtaactat cgtcttgagt 4740
ccaacccggg aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4800
gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca 4860
ctagaagaac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4920
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtaggtttt tttgtttgca 4980
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5040
ggctctgacgc tcagtggaaac gaaaactcac gttaagggat tttgggtcatg agattatcaa 5100
aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatac atctaaagta 5160
tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 5220
cgatctgtct atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc 5280
tgctcgtga agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc 5340
agaaagtga ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5400
tgaacttttg ctttgccacg gaacggctcg cgttgtcggg aagatgcgtg atctgatcct 5460
tcaactcagc aaaagttcga tttattcaac aaagccgccg tcccgtaag tcagcgtaat 5520
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5580
atgaaactgc aattttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5640
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5700
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttccccct cgtcaaaaat 5760
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag 5820
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 5880
actcgcatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 5940
atcgctgtta aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc 6000
cagcgcatca acaatatatt cacctgaatc aggatattct tctaatacct ggaatgctgt 6060
tttcccgggg atcgcagtgg tgagtaacca tgcacatca ggagtaacgga taaaatgctt 6120
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6180
atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6240
atacaatcga tagattgtcg cacctgattg ccgcacatta tcgcgagccc atttataccc 6300
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6360
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6420
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6480
ttcccccccc cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6540
atlttgatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6600
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 6660
cacgaggccc tttcgtc

```

&lt;210&gt; 121

&lt;211&gt; 6677

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 121

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg ttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctttttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgtctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcact ggaagcagt gtggatctgg ggcacctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagcca ccctgttctg cgccagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gacccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gacctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgatcgatc aactgcaaca 2400
ccagcgccat caccaggcc tgcaccaagg tgagcttcga ccccatcccc atccactact 2460
gcgccccgc cggctacgcc atcctgaagt gcaacaacaa gaccttcaac ggcaccggcc 2520
cctgcaacaa cgtgagcacc atccagtgc cccacggcac caagcccgtg gtgagcacc 2580
agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc aagaacctga 2640
ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcacc 2700
gccgatcggc cactgcaac atcagccgca ccgcctggaa caagaccctg caggaggtgg 2760
gcaagaagct ggccgagcac tttcccaaca aggccatcaa gttcgccaag cacagcggcg 2820
gcgacctgga gatcaccacc cacagcttca actgccgcgg cgagtctctc tactgcaaca 2880
ccagcagcct gttcaacagc acctacacc ccaacagcac cgagaacatc accggcaccg 2940
agaacagcat catcaccatc ccctgccgca tcaagcagat catcaacatg tggcagggcg 3000
tgggcgcgc catgtacgcc cccccatcg agggcatcct gacctgccgc agcaacatca 3060
ccggcctgct gctgacccgc gacggcgga cggcatgca cgacaccgag atcttccgcc 3120
ccgagggcg cgacatgcgc gacaactggc gcagcgagct gtacaagtac aagggtggtg 3180

```



```

agatcaagcc cctgggcatc gccccacca aggccaaagct taccgtgcag gtgcgccagc 3240
tgctgagcgg catcgtgcag cagcagagca acctgctgcg cgccatcgag gccagcagc 3300
acatgctgca gctgaccgtg tggggcatca agcagctgca gaccgcgctg ctggccatcg 3360
agcgtacct gcgcgaccag cagctcgagg acatctggaa caacatgacc tggatgcagt 3420
gggaccgcga gatcagcaac tacaccaaca ccatctaccg cctgctggag gacagccaga 3480
accagcagga gaagaacgag caggacctgc tggccctgga caagtggcag aacctgtgga 3540
cctgggtcgg catcaccaac tggctgtggt acatcaagat cttcatcaag atcgtgggcg 3600
gcctgatcgg cctgcgcac atcttcgccc tgctgagcat ctgaacacgt gggatccaga 3660
tctgctgtgc cttctagttg ccagccatct gttgtttgcc cctccccgt gccttccttg 3720
acctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat 3780
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 3840
gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggg taccaggtg 3900
ctgaagaatt gaccgggttc ctccctgggc agaaagaagc aggcacatcc ctttctctgt 3960
gacacacct gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag 4020
ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggg ctctccctcc 4080
ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 4140
ggctattaag tgcagagggg gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 4200
tcatagaatt ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttcctcgc 4260
tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4320
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4380
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggetcc 4440
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagc 4500
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttcgga 4560
ccctgccgt taccggatc ctgtccgcct ttctcccttc ggaagcgtg gcgctttctc 4620
atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 4680
tgacgaacc ccccgttcag cccgaccgct gcgccttata cggtaactat cgtcttgagt 4740
ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4800
gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca 4860
ctagaagaac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4920
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca 4980
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 5040
ggtctgacgc tcagtggaa gaaaactcac gttaagggat tttggtcatg agattatcaa 5100
aaaggatctt cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta 5160
tatatgagta aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag 5220
cgatctgtct atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc 5280
tgccctcgtga agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc 5340
agaaagtgag ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt 5400
tgaacttttg ctttgccacg gaacgggtctg cgttgtcggg aagatgcgtg atctgatcct 5460
tcaactcagc aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat 5520
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5580
atgaaactgc aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt 5640
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg 5700
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 5760
aaggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag 5820
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 5880
actcgcatca accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg 5940
atcgctgtta aaaggacaat tacaacagg aatcgaatgc aaccggcgca ggaacactgc 6000
cagcgcatca acaatathtt cacctgaatc aggatattct tctaatacct ggaatgctgt 6060
tttcccgggg atcgcagtg tgagtaacca tgcacatca ggagtacgga taaaatgctt 6120
gatggtcgga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 6180
atcatgggca acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc 6240
atacaatcga tagattgtcg cacctgattg cccgacatta tgcgagccc atttataccc 6300
atataaatca gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg 6360
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6420
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6480

```



```

ttcccccccc cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6540
at ttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6600
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgat 6660
cacgaggccc ttctgtc 6677

```

&lt;210&gt; 122

&lt;211&gt; 6608

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 122

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcgtt ttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttgc gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttgc gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccattttatta tttaaaaatt cacatataca acaacgcgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggtctt tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcgggtcat ggtcgtcgtg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagt gtggatctgg ggcaccttg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagcca ccctgttctg cgcagcgac gccaaaggct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340

```

gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	ccgatcgatc	aactgcaaca	2400
ccagecgccat	caccagggcc	tgccccaagg	tgagcttcga	cccatcccc	atccactact	2460
gcgccccgc	cggctacgcc	atcctgaagt	gcaacaacaa	gaccttcaac	ggcaccggcc	2520
cctgcaacaa	cgtgagcacc	atccagtgc	cccacggcac	caagcccgtg	gtgagcacc	2580
agctgctgct	gaacggcagc	ctggccgagg	aggagatcat	catccgcagc	aagaacctga	2640
ccgacaacgt	gaagaccatc	atcgtgcacc	tgaacgagag	cgtggagatc	aactgcaccc	2700
gccgatcggc	ccactgcaac	atcagccgca	ccgcctggaa	caagaccctg	caggaggtgg	2760
gcaagaagct	ggccgagcac	ttccccaaca	aggccatcaa	gttcgccaag	cacagcggcg	2820
gcgacctgga	gatcaccacc	cacagcttca	actgccgcgg	cgagttcttc	tactgcaaca	2880
cccgatcgat	cccctgccgc	atcaagcaga	tcataacat	gtggcagggc	gtgggcccgc	2940
ccatgtacgc	ccccccatc	gagggcatcc	tgacctgccg	cagcaacatc	accggcctgc	3000
tgctgaccgc	cgacggcggc	accggcatgc	acgacaccga	gatcttccgc	cccaggggcg	3060
gcgacatgcg	cgacaactgg	cgcagcgagc	tgtacaagta	caaggtgggtg	gagatcaagc	3120
ccctgggcat	cgccccacc	aaggccaagc	ttaccgtgca	ggtgcgccag	ctgctgagcg	3180
gcatcgtgca	gcagcagagc	aacctgctgc	gcgccatcga	ggcccagcag	cacatgctgc	3240
agctgaccgt	gtggggcatc	aagcagctgc	agaccgcgt	gctggccatc	gagcgctacc	3300
tgcgcgacca	gcagctcgag	gacatctgga	acaacatgac	ctggatgcag	tgggaccgcg	3360
agatcagcaa	ctacaccaac	accatctacc	gcctgctgga	ggacagccag	aaccagcagg	3420
agaagaacga	gcaggacctg	ctggccctgg	acaagtggca	gaacctgtgg	acctggctcg	3480
gcatcaccaa	ctggctgtgg	tacatcaaga	tcttcatcaa	gatcgtgggc	ggcctgatcg	3540
gcctgcgcat	catcttcgcc	gtgctgagca	tctgaacacg	tgggatccag	atctgctgtg	3600
ccttctagtt	gccagccatc	tggtgtttgc	ccctcccccg	tgcttctctt	gaccttgga	3660
ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	3720
aggtgtcatt	ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	3780
gacaatagca	ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccaggt	gctgaagaat	3840
tgacccggtt	cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	3900
tgtccacgcc	cctgggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	3960
gctccgcctt	caatcccacc	cgctaaagta	cttgagcgcg	tcttccctc	cctcatcagc	4020
ccaccaaac	aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	4080
gtgcagaggg	agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	4140
tttaaggcca	tgatttaagg	ccatcatggc	cttaatcttc	cgttccctcg	ctcactgact	4200
cgtgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatat	4260
ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	4320
aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgcctccctg	4380
acgagcatca	caaaaatcga	cgtcaagtc	agaggtggcg	aaaccgcaca	ggactataaa	4440
gataccaggc	gtttccctct	ggaagctccc	tcgtagcgctc	tcctgttccg	acctgcccgc	4500
ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	4560
gctgtaggta	tctcagttcg	gtgtaggctg	ttcgctccaa	gctgggctgt	gtgcacgaac	4620
ccccggttca	gcccagccgc	tgcgcttat	ccggtaaacta	tcgtcttgag	tccaaccg	4680
taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	4740
atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagaa	4800
cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggtagct	4860
cttgatccgg	caaacaaacc	accgctggta	gcggtgggtt	ttttgtttgc	aagcagcaga	4920
ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	4980
ctcagtggaa	cgaaaactca	cgttaaggga	ttttgggtcat	gagattatca	aaaaggatct	5040
tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	aatctaaagt	atatatgagt	5100
aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	acctatctca	gcgatctgtc	5160
tatttctgtt	atccatagtt	gcctgactcg	gggggggggg	gcgctgaggt	ctgcctcgtg	5220
aagaagggtg	tgctgactca	taccaggcct	gaatcgcccc	atcatccagc	cagaaagtga	5280
gggagccacg	gttgatgaga	gctttgttgt	aggtggacca	gttggtgatt	ttgaactttt	5340
gctttgccac	ggaacgggtc	gcgttgctcg	gaagatgcgt	gatctgatcc	ttcaactcag	5400
caaaagtctg	atbtattcaa	caaagccgcc	gtcccgtcaa	gtcagcgtaa	tgctctgcca	5460
gtgttacaac	caattaacca	attctgatta	gaaaaactca	tcgagcatca	aatgaaactg	5520
caatttatct	atattcaggat	tatcaatacc	atatttttga	aaaagccgtt	tctgtaatga	5580
aggagaaaac	tcaccgaggc	agttccatag	gatggcaaga	tcctgggtatc	ggtctgcgat	5640

```

tccgactcgt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc 5700
aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat 5760
ttctttccag acttggtcaa caggccagcc attacgctcg tcatcaaaat cactcgcac 5820
aaccaaaccg ttattcattc gtgattgctc ctgagcgaga cgaaatacgc gatcgctgtt 5880
aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac 5940
aacaatatat tcacctgaat caggatattc ttctaatacc tggaatgctg ttttcccggg 6000
gatcgagtg gtgagtaacc atgcatcac aggagtacgg ataaaatgct tgatggtcgg 6060
aagaggcata aattccgtca gccagtttag tctgaccac tcatctgtaa catcattggc 6120
aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg 6180
atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc 6240
agcatccatg ttggaattta atcgcgccct cgagcaagac gtttcccgtt gaatatggct 6300
cataacaccc cttgtattac tgtttatgta agcagacagt tttattgttc atgatgat 6360
atthtttatct tgtgcaatgt aacatcagag atthttgagac acaacgtggc tttccccccc 6420
ccccattat tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg 6480
tatttagaaa aataaaca aa taggggttcc ggcacattt ccccgaaaag tgccacctga 6540
cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc 6600
ctttcgtc 6608

```

&lt;210&gt; 123

&lt;211&gt; 6689

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 123

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gtccgcgctt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aacgacccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacac aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccga ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatc tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatatata acaacgcgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggctcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560

```

agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920  
ccatgcgcgt ggagggcacc cagcgcaact ggaagcagtg gtggatctgg ggcatcctgg 1980  
gcttctggat ggtgatgac tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040  
gcgtgcccgt gtggaaggag gccaaagacca ccctgttctg cggcagcgac gccaaaggcct 2100  
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160  
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220  
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280  
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340  
gcaccgacaa catgggcggc gagatcaaga actgcagctt ccgacgatc aactgcaaca 2400  
ccagcgccat caccaggcc tgcaccaagg tgagcttcca ccccatcccc atccactact 2460  
gcgccccgc cggctacgcc atcctgaagt gcaacaacaa gaccttcaac ggcaccggcc 2520  
cctgcaacaa cgtgagcacc atccagtga cccacggcac caagcccgtg gtgagcacc 2580  
agctgctgct gaacggcagc ctggccgagg aggagatcat catccgcagc aagaacctga 2640  
ccgacaacgt gaagaccatc atcgtgcacc tgaacgagag cgtggagatc aactgcacc 2700  
gccccaaacaa caacaccgc aagagcatcc gcatcgcccc cggccaggcc ttctacgcca 2760  
ccggcgagat catcggcgac atccgccagg cccactgcaa catcagccgc accgcctgga 2820  
acaagaccct gcaggagggt ggcaagaagc tggccgagca cttccccaac aaggccatca 2880  
agttcgccaa gcacagcgcc ggcgacctgg agatcaccac ccacagcttc aactgccgcg 2940  
gcgagttctt ctactgcaac acccgatcga tccctgccc catcaagcag atcatcaaca 3000  
tgtggcaggg cgtgggcgc gccatgtacg cccccccat cgagggcac ctgacctgcc 3060  
gcagcaacat caccggcctg ctgctgaccc gcgacggcgg caccggcatg cagcaccgc 3120  
agatcttccg ccccgagggc ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt 3180  
acaaggtggt ggagatcaag cccctgggca tcgccccac caaggccaag cttaccgtgc 3240  
agggtgcgca gctgctgagc ggcacgtgc agcagcagag caacctgctg cgcgccatcg 3300  
aggcccagca gcacatgctg cagctgaccg tgtggggcat caagcagctg cagacccgcg 3360  
tgctggccat cgagcgctac ctgcgcgacc agcagctcga ggacatctgg aacaacatga 3420  
cctggatgca gtgggaccgc gagatcagca actacaccaa caccatctac cgcctgctgg 3480  
aggacagcca gaaccagcag gagaagaacg agcaggacct gctggccctg gacaagtggc 3540  
agaacctgtg gacctgggtc ggcacacca actggctgtg gtacatcaag atcttcatca 3600  
agatcgtggg cggcctgatc ggcctgcgca tcatcttcgc cgtgctgagc atctgaacac 3660  
gtgggatcca gatctgctgt gccttctagt tgccagccat ctgttgtttg cccctcccc 3720  
gtgccttcc taccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa 3780  
attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac 3840  
agcaaggggg aggatggga agacaatagc aggcacgtcg gggatgcggt gggctctatg 3900  
ggtacccagg tgctgaagaa ttgaccgggt tctcctggg ccagaaagaa gcaggcacat 3960  
ccccttctct gtgacacacc ctgtccacgc ccctggttct tagttccagc cccactcata 4020  
ggacactcat agctcaggag ggctccgcct tcaatcccac ccgctaaagt acttgagcg 4080  
gtctctccct ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaagaaat 4140  
taaagcaaga taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag 4200  
taatgagaga aatcatagaa ttttaaggcc atgatttaag gccatcatgg ccttaatctt 4260  
ccgcttccct gctcactgac tcgctgcgct cggctcgttcg gctgcggcga gcggtatcag 4320  
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 4380  
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcggtg ctggcgtttt 4440  
tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagagggtggc 4500  
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctgctgcgct 4560  
ctcctgttcc gacctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 4620  
tggcgcttct tcatagctca cgtgttaggt atctcagttc ggtgtaggtc gttcgctcca 4680  
agctgggctg tgtgcacgaa cccccgttc agcccagccg ctgcgcctta tccggtaact 4740  
atcgtcttga gtccaacccg gtaagacagc acttatcgcc actggcagca gccactggta 4800  
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta 4860



```

actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct 4920
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggg agcgggtggt 4980
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 5040
tcttttctac ggggtctgac gctcagtggg acgaaaactc acgttaaggg attttgggtca 5100
tgagattatc aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agttttaaat 5160
caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg 5220
cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc gggggggggg 5280
ggcgctgagg tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc 5340
catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 5400
agttgggtgat tttgaacttt tgctttgcca cggaacgggt tgcgttgctg ggaagatgcg 5460
tgatctgate cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca 5520
agtcagcgta atgctctgcc agtggtacaa ccaattaacc aattctgatt agaaaaactc 5580
atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 5640
aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 5700
atcctgggtat cgggtctgcg ttccgactcg tccaacatca atacaacct ttaatttccc 5760
ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccgggtga 5820
gaatggcaaaa agcttatgca tttctttcca gacttggtca acaggccagc cattacgctc 5880
gtcatcaaaa tcactcgcac caaccaaacc gttattcatt cgtgattgct cctgagcgag 5940
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg 6000
caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac 6060
ctggaatgct gttttcccg ggatcgcat ggtgagtaac catgcatcat caggagtacg 6120
gataaaatgc ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgaccat 6180
ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc 6240
atcgggcttc ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc 6300
ccatttatac ccatataaat cagcatccat gttggaattt aatcgcgcc tcgagcaaga 6360
cgtttcccg tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag 6420
ttttattggt catgatgata tatttttata ttgtgcaatg taacatcaga gattttgaga 6480
cacaacgtgg ctttcccccc cccccatta ttgaagcatt tatcagggtt attgtctcat 6540
gagcggatag atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt 6600
tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa 6660
aaataggcgt atcacgaggc cctttcgtc 6689

```

&lt;210&gt; 124

&lt;211&gt; 6785

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 124

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720

```



catcaatggg	cgtggatagc	ggtttgactc	acgggggattt	ccaagtctcc	accccatgga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctataactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggtc	ccattttatta	tttacaattt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgccgggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgcgt	ggagggcatc	cagcgcgaact	ggaagcagtg	gtggatctgg	ggcatcctgg	1980
gcttctggat	ggtgatgac	tacaacgtgc	gcggcaacct	gtgggtgacc	gtgtactacg	2040
gcgtgcccgt	gtggaaggag	gccaagacca	ccctgttctg	cgccagcgac	gccaaggcct	2100
acgacgccga	ggtgcacaac	gtgtgggcca	cccacgcctg	cgtgcccacc	gaccccaacc	2160
cccaggagat	ggtgctggag	aacgtgaccg	agaacttcaa	catgtgggag	aacgacatgg	2220
tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	caacatgacc	accgaggtgc	2400
gcgacaagcg	cgagaagggtg	cacgcctgt	tctaccgctt	ggacatcggtg	cccctgaaga	2460
acgagagcag	caacaccagc	ggcgactacc	gcctgatcaa	ctgcaacacc	agcgccatca	2520
cccaggcctg	ccccaaagggtg	agcttcgacc	ccatccccat	ccactactgc	gccccgcgg	2580
gctacgccat	cctgaagtgc	aacaacaaga	ccttcaacgg	caccggcccc	tgcaacaacg	2640
tgagcaccat	ccagtgcacc	cacggcacca	agcccggtgg	gagcaccag	ctgctgctga	2700
acggcagcct	ggccgaggag	gagatcatca	tccgcagcaa	gaacctgacc	gacaacgtga	2760
agaccatcat	cgtgcacctg	aacgagagcg	tggagatcaa	ctgcacccgc	cgatcggccc	2820
actgcaacat	cagccgcacc	gcctggaaca	agaccctgca	ggaggtgggc	aagaagctgg	2880
ccgagcactt	ccccaaacaag	gccatcaagt	tcgccaagca	cagcggcggc	gacctggaga	2940
tcaccaccca	cagcttcaac	tgcgcggcg	agttcttcta	ctgcaacacc	agcagcctgt	3000
tcaacagcac	ctacaccccc	aacagcaccg	agaacatcac	cggcaccgag	aacagcatca	3060
tcaccatccc	ctgccgcac	aagcagatca	tcaacatgtg	gcagggcggtg	ggccgcgcca	3120
tgtacgcccc	ccccatcgag	ggcatcctga	cctgccgcag	caacatcacc	ggcctgctgc	3180
tgaccgcgga	cggcggcacc	ggcatgcacg	acaccgagat	cttcgcgcgc	gagggcgggc	3240
acatgcgcga	caactggcgc	agcgagctgt	acaagtacaa	gggtggtggag	atcaagcccc	3300
tgggcatcgc	ccccaccaag	gccaagctta	ccgtgcaggt	gcgccagctg	ctgagcggca	3360
tcgtgcagca	gcagagcaac	ctgctgcgcg	ccatcgaggc	ccagcagcac	atgctgcagc	3420
tgaccgtgtg	gggcatcaag	cagctgcaga	cccgcgtgct	ggccatcgag	cgctacctgc	3480
gcgaccagca	gctcgaggac	atctggaaca	acatgacctg	gatgcagtg	gaccgcgaga	3540
tcagcaacta	caccaacacc	atctaccgcc	tgttgaggga	cagccagaac	cagcaggaga	3600
agaacgagca	ggacctgctg	gccctggaca	agtggcagaa	cctgtggacc	tgggtcggca	3660
tcaccaactg	gctgtggtac	atcaagatct	tcatcaagat	cgtgggcggc	ctgatcggcc	3720
tgcgcacat	cttcgcgcgtg	ctgagcatct	gaacacgtgg	gatccagatc	tgtgtgcct	3780
tctagttgcc	agccatctgt	tgtttgcccc	ttccccgtgc	cttccttgac	cctggaagg	3840
gccactccca	ctgtccttcc	ctaataaaat	gaggaaattg	catcgcatg	tctgagtagg	3900
tgtcattcta	ttctgggggg	tgggggtggg	caggacagca	agggggagga	ttgggaagac	3960
aatagcaggc	atgctggggg	tgcgggtggg	tctatgggta	cccagggtgct	gaagaattga	4020

```

ccccggttcct cctgggcccag aaagaagcag gcacatcccc ttctctgtga cacaccctgt 4080
ccacgccccct ggttccttagt tccagcccca ctcataggac actcatagct caggagggct 4140
ccgccttcaa tcccacccgc taaagtactt ggagcggctt ctccctccct catcagccca 4200
ccaaacaaaa cctagcctcc aagagtggga agaaattaaa gcaagatagg ctattaagtg 4260
cagaggggaga gaaaatgcct ccaacatgtg aggaagtaat gagagaaatc atagaatttt 4320
aaggccatga ttttaaggcca tcatggcctt aatcttccgc ttctctgctc actgactcgc 4380
tgcgctcggg cgttcgggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg 4440
tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaaggc cagcaaaaagg 4500
ccaggaaccg taaaaaggcc gcgttgctgg cgttttttcca taggctccgc cccctgacg 4560
agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 4620
accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 4680
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 4740
gtaggtatct cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 4800
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtta 4860
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 4920
taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag 4980
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 5040
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 5100
cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 5160
agtgaacga aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca 5220
cctagatcct tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa 5280
cttggtctga cagttaccaaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat 5340
ttcgttcata catagttgcc tgactcgggg gggggggggcg ctgaggtctg cctcgtgaag 5400
aagggtgttg tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg 5460
agccacgggt gatgagagct ttgttgtagg tggaccagtt ggtgattttg aacttttgct 5520
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgacccctc aactcagcaa 5580
aagttcgatt tattcaacaa agccgcctgc ccgtcaagtc agcgtaatgc tctgccagtg 5640
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 5700
tttattcata tcaggattat caataccata tttttgaaaa agccgtttct gtaatgaagg 5760
agaaaactca ccgaggcagt tccataggat ggcaagatcc tggatcgggt ctgcgattcc 5820
gactcgtcca acatcaatac aacctattaa tttccctcgc tcaaaaataa ggttatcaag 5880
tgagaaatca ccatgagtga cgactgaatc cggtgagaat ggcaaaaagct tatgcatttc 5940
tttccagact tgttcaacag gccagccatt acgctcgtca tcaaaatcac tcgcatcaac 6000
caaaccgtta ttcattcgtg attgcgcctg agcgagacga aatacgcgat cgctgttaaa 6060
aggacaatta caaacaggaa tcgaatgcaa ccggcgagg aacactgcca gcgcatcaac 6120
aatattttca cctgaatcag gatattcttc taatacctgg aatgctgttt tcccggggat 6180
cgagtggtg agtaaccatg catcatcagg agtacggata aaatgcttga tggtcggaag 6240
aggcataaat tccgtcagcc agtttagtct gaccatctca tctgtaacat cattggcaac 6300
gctacctttg ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata 6360
gattgtcgca cctgattgcc cgacattatc gcgagcccat ttatacccat ataaatcagc 6420
atccatgttg gaatttaatc gcggcctcga gcaagacgtt tcccgttgaa tatggctcat 6480
aacacccctt gtattactgt ttatgtaagc agacagtttt attgttcatt atgatattt 6540
tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccccccccc 6600
ccattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 6660
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt 6720
ctaagaaacc attattatca tgacattaac ctataaaaaat aggcgtatca cgaggccctt 6780
tcgtc 6785

```

&lt;210&gt; 125

&lt;211&gt; 6716

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 125

```
tgcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattgg 420
cccgcctggc tgaccgccc aacgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctggt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgaggct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgcgg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacccctg 1980
gcttctggat ggtgatgata tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagcca ccctgttctg cgccagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca cccacgcctg cgtgcccacc gaccccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220
tggagcagat gcaccaggac atcatcagcc tgtgggacca gagcctgaag ccctgcgtga 2280
agctgacccc cctgtgcgtg accctgcact gcagcaaccg caccatcgac tacaacaacc 2340
gcaccgacaa catgggcggc gagatcaaga actgcagctt caacatgacc accgaggtgc 2400
gcgacaagcg cgagaagggt cagcctctgt tctaccgcct ggacatcgtg cccctgaaga 2460
acgagagcag caacaccagc ggcgactacc gcctgatcaa ctgcaacacc agcgccatca 2520
cccaggcctg ccccaagggt agcttcgacc ccatccccat ccactactgc gcccccgccg 2580
gctacgccat cctgaagtgc aacaacaaga ccttcaacgg caccggcccc tgcaacaacg 2640
tgagcaccat ccagtgcacc cagggacca agcccgtggt gagcaccag ctgctgctga 2700
acggcagcct ggccgaggag gagatcatca tccgcagcaa gaacctgacc gacaacgtga 2760
agaccatcat cgtgcacctg aacgagagcg tggagatcaa ctgcaccgc cgatcgccc 2820
actgcaacat cagccgcacc gcctggaaca agaccctgca ggagggtggg aagaagctgg 2880
ccgagcactt cccaacaag gccatcaagt tcgccaagca cagcggcggc gacctggaga 2940
tcaccacca cagcttcaac tgccgcgcg agttcttcta ctgcaacacc cgatcgatcc 3000
cctgcccgat caagcagatc atcaacatgt ggcaggcggt gggccgcgcc atgtacgccc 3060
```



```

cccccatcga gggcatcctg acctgccgca gcaacatcac cggcctgctg ctgaccccgcg 3120
acggcggcac cggcatgcac gacaccgaga tcttccgccc cgagggcggc gacatgcgcg 3180
acaactggcg cagcgagctg tacaagtaca aggtggtgga gatcaagccc ctgggcatcg 3240
ccccaccaa ggccaagctt accgtgcagg tgcgccagct gctgagcggc atcgtgcagc 3300
agcagagcaa cctgctgcgc gccatcgagg ccagcagca catgctgcag ctgaccgtgt 3360
ggggcatcaa gcagctgcag acccgctgc tggccatcga gcgctacctg cgcgaccagc 3420
agctcgagga catctggaac aacatgacct ggatgcagtg ggaccgcgag atcagcaact 3480
acaccaacac catctaccgc ctgctggagg acagccagaa ccagcaggag aagaacgagc 3540
aggacctgct ggccctggac aagtggcaga acctgtggac ctggttcggc atcaccaact 3600
ggctgtggta catcaagatc ttcatcaaga tcgtgggcgg cctgatcggc ctgcgcatca 3660
tcttcgccgt gctgagcatc tgaacacgtg ggatccagat ctgctgtgcc ttctagttgc 3720
cagccatctg ttgtttgccc ctcccccgct ccttccttga ccctggaagg tgccactccc 3780
actgtccttt cctaataaaa tgaggaaatt gcatcgcat gtctgagtag gtgtcattct 3840
attctggggg gtgggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg 3900
catgctgggg atgcgggtggg ctctatgggt acccaggtgc tgaagaattg acccggttcc 3960
tcctgggcca gaaagaagca ggcacatccc cttctctgtg acacaccctg tccacgcccc 4020
tggttcttag ttccagcccc actcatagga cactcatagc tcaggagggc tccgccttca 4080
atcccacccg ctaaagtact tggagcggtc tctccctccc tcatcagccc accaaaccaa 4140
acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag 4200
agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg 4260
atttaaggcc atcatggcct taatcttccg cttcctcgct cactgactcg ctgcgctcgg 4320
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 4380
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 4440
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca 4500
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 4560
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 4620
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 4680
tcagttcggg gtaggtcggt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc 4740
ccgaccgctg cgcttatcc ggtaactatc gtcttgagtc caaccggta agacacgact 4800
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcggtg 4860
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg 4920
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 4980
aacaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 5040
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 5100
aaaactcacg ttaagggtatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 5160
ttttaaatba aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttgggtctg 5220
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta ttctgttcat 5280
ccatagttgc ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaagggtgtg 5340
ctgactcata ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt 5400
tgatgagagc tttggtgtag gtggaccagt tgggtathtt gaacttttgc tttgccacgg 5460
aacggtctgc gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat 5520
ttattcaaca aagccgcgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca 5580
attaaccaat tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat 5640
atcaggatta tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc 5700
accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc 5760
aacatcaata caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc 5820
accatgagtg acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac 5880
ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctgcacatcaa ccaaaccggt 5940
attcatctcg gattgcgcct gagcgagacg aaatacgcg aagcgcacaa caatatcttc 6000
acaaacagga atcgaatgca accggcgcg gaacactgcc agcgcaccaa caatatcttc 6060
acctgaatca ggatattctt ctaataacct gaatgctgtt ttcccgggga tcgcagtggg 6120
gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa 6180
ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt 6240
gccatgtttc agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc 6300
acctgattgc ccgacattat cgcgagccca ttatatacca tataaatcag catccatggt 6360

```

```

ggaattttaat cgcggcctcg agcaagacgt ttcccgttga atatggctca taacacccct 6420
tgtattactg tttatgtaag cagacagttt tattgttcat gatgatabat ttttatcttg 6480
tgcaatgtaa catcagagat tttgagacac aacgtggctt tccccccccc cccattattg 6540
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6600
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 6660
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6716

```

&lt;210&gt; 126

&lt;211&gt; 6797

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 126

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgcgt ggagggcatc cagcgcaact ggaagcagtg gtggatctgg ggcacctgg 1980
gcttctggat ggtgatgatc tacaacgtgc gcggcaacct gtgggtgacc gtgtactacg 2040
gcgtgcccgt gtggaaggag gccaaagaca ccctgttctg cgcagcgac gccaaaggcct 2100
acgacgccga ggtgcacaac gtgtgggcca ccacgcctg cgtgcccacc gacccaacc 2160
cccaggagat ggtgctggag aacgtgaccg agaacttcaa catgtgggag aacgacatgg 2220

```



tggagcagat	gcaccaggac	atcatcagcc	tgtgggacca	gagcctgaag	ccctgcgtga	2280
agctgacccc	cctgtgcgtg	accctgcact	gcagcaaccg	caccatcgac	tacaacaacc	2340
gcaccgacaa	catgggcggc	gagatcaaga	actgcagctt	caacatgacc	accgaggtgc	2400
gcgacaagcg	cgagaagggtg	cacgccctgt	tctaccgcct	ggacatcgtg	cccctgaaga	2460
acgagagcag	caacaccagc	ggcgactacc	gcctgatcaa	ctgcaacacc	agcgccatca	2520
cccaggcctg	ccccaagggtg	agcttcgacc	ccatccccc	ccactactgc	gccccgcgcg	2580
gctacgccat	cctgaagtgc	aacaacaaga	ccttcaacgg	caccggcccc	tgcaacaacg	2640
tgagcaccat	ccagtgcacc	cacggcacca	agcccgtggt	gagcaccag	ctgctgctga	2700
acggcagcct	ggccgaggag	gagatcatca	tccgcagcaa	gaacctgacc	gacaacgtga	2760
agaccatcat	cgtgcacctg	aacgagagcg	tggagatcaa	ctgcaccgcg	cccaacaaca	2820
acacccgcaa	gagcatccgc	atcgggccccg	gccaggcctt	ctacgccacc	ggcgagatca	2880
tcggcgacat	ccgccaggcc	cactgcaaca	tcagccgcac	cgcctggaac	aagaccctgc	2940
aggaggtggg	caagaagctg	gccgagcact	tccccaacaa	ggccatcaag	ttcgccaagc	3000
acagcggcgg	cgacctggag	atcaccaccc	acagcttcaa	ctgccgcggc	gagttcttct	3060
actgcaacac	ccgatcgatc	ccctgccgca	tcaagcagat	catcaacatg	tggcagggcg	3120
tgggcccgcg	catgtacgcc	ccccccatcg	agggcatcct	gacctgccgc	agcaacatca	3180
ccggcctgct	gctgacccgc	gacggcggca	ccggcatgca	cgacaccgag	atcttccgcc	3240
ccgagggcgg	cgacatgcgc	gacaactggc	gcagcgagct	gtacaagtac	aagggtggtg	3300
agatcaagcc	cctgggcatc	gccccacca	aggccaagct	taccgtgcag	gtgcgccagc	3360
tgctgagcgg	catcgtgcag	cagcagagca	acctgctgcg	cgccatcgag	gcccagcagc	3420
acatgctgca	gctgaccgtg	tggggcatca	agcagctgca	gacccgcgtg	ctggccatcg	3480
agcgctacct	gcgcgaccag	cagctcgagg	acatctggaa	caacatgacc	tggatgcagt	3540
gggaccgcga	gatcagcaac	tacaccaaca	ccatctaccg	cctgctggag	gacagccaga	3600
accagcagga	gaagaacgag	caggacctgc	tggccctgga	caagtggcag	aacctgtgga	3660
cctggttcgg	catcaccaac	tggctgtggt	acatcaagat	cttcatcaag	atcgtgggcg	3720
gcctgatcgg	cctgcgcac	atcttcgcgc	tgctgagcat	ctgaacacgt	gggatccaga	3780
tctgctgtgc	cttctagtgt	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	3840
accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcatcgcat	3900
tgtctgagta	ggtgtcattc	tattctgggg	ggtgggggtg	ggcaggacag	caagggggag	3960
gattgggaag	acaatagcag	gcattgctgg	gatgcgggtg	gctctatggg	taccaggtg	4020
ctgaagaatt	gacccggttc	ctcctgggccc	agaaagaagc	aggcacatcc	ccttctctgt	4080
gacacaccc	gtccacgccc	ctggttctta	gttccagccc	cactcatagg	acactcatag	4140
ctcaggagg	ctccgccttc	aatcccaccc	gctaaagtac	ttggagcgg	ctctccctcc	4200
ctcatcagcc	caccaaacca	aacctagcct	ccaagagtgg	gaagaaatta	aagcaagata	4260
ggctattaag	tgagaggga	gagaaaatgc	ctccaacatg	tgaggaaagta	atgagagaaa	4320
tcatagaatt	ttaaggccat	gattttaagg	catcatggcc	ttaatcttcc	gcttcctcgc	4380
tcactgactc	gctgcgctcg	gtcgttcggc	tgccggcagc	ggtatcagct	cactcaaagg	4440
cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	4500
gccagcaaaa	ggccagggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	4560
gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	4620
gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	4680
ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	4740
atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	4800
tgcacgaacc	ccccgttcag	cccagaccgt	gcgccttatc	cggtaactat	cgtcttgagt	4860
ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	4920
gagcgaggta	tgtaggcgg	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	4980
ctagaagaac	agtatttggt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	5040
ttggtagctc	ttgatccggc	aaacaaaacca	ccgctggtag	cgggtggttt	tttgtttgca	5100
agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	5160
ggtctgacgc	tcagtggaa	gaaaactcac	gttaagggat	tttgggtcatg	agattatcaa	5220
aaaggatctt	cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	5280
tatatgagta	aacttgggtc	gacagttacc	aatgcttaab	cagtgaggca	cctatctcag	5340
cgatctgtct	atttcgttca	tccatagttg	cctgactcgg	gggggggggg	cgctgaggtc	5400
tgccctcgtga	agaagggtgt	gctgactcat	accaggcctg	aatcgcccca	tcattccagcc	5460
agaaagtga	ggagccacgg	ttgatgagag	ctttgttgta	ggtggaccag	ttggtgattt	5520

```

tgaacttttg ctttgccacg gaacggtctg cgttggtcggg aagatgcgtg atctgaccc 5580
tcaactcagc aaaagttcga tttattcaac aaagccgccc tcccgtcaag tcagcgtaat 5640
gctctgccag tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa 5700
atgaaactgc aattttattca tatcaggatt atcaatacca tttttttgaa aaagccggtt 5760
ctgtaatgaa ggagaaaact caccgaggca gttccatagg atggcaagat cctgggatcg 5820
gtctgcgatt ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat 5880
aagggttatca agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag 5940
cttatgcatt tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc 6000
actcgcatca accaaaccgt tattcattcg tgattgcgce tgagcgagac gaaatacgcg 6060
atcgctgtta aaaggacaat tacaacagg aatcgaatgc aaccggcgca ggaacactgc 6120
cagcgcatca acaatatattt cacctgaatc aggatattct tctaatacct ggaatgctgt 6180
tttcccgggg atcgcgatgg tgagtaacca tgcacatca ggagtacgga taaaatgctt 6240
gatggtcggg agaggcataa attccgctcag ccagtttagt ctgacctct catctgtaac 6300
atcattggca acgctacctt tgccatggtt cagaaacaac tctggcgcat cgggcttccc 6360
atacaatcga tagattgtcg cacctgattg cccgacatta tcccgagccc atttataccc 6420
atataaatca gcatccatgt tggaatttaa tccgggcctc gagcaagacg tttcccgttg 6480
aatatggctc ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca 6540
tgatgatata tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct 6600
ttcccccccc cccattatt gaagcattta tcagggttat tgtctcatga gcggatacat 6660
atgtgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 6720
gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcggtat 6780
cacgaggccc tttcgtc

```

&lt;210&gt; 127

&lt;211&gt; 6718

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 127

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatggtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aacgacccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgagggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaaaaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260

```

tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttataaaatt	cacatatata	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcccttccat	gggtcttttc	1860
tgcagtccac	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tctgtggatac	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggatccctgg	1980
gtttctggat	gatcatcatc	tgcggtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggg	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacac	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagccgtaac	tgtccttca	2340
acaccaccac	cgaaatccgt	gacaaaaaac	agcagggtta	cgtctgttc	taccgtccgg	2400
acatcgttct	gctgaaagaa	aaccgtaaca	actccaacaa	ctccgaatac	atcctgatca	2460
actgcaacgc	ttccaccatc	accaggett	gcccgaaggt	taacttcgac	ccgatcccga	2520
tccactactg	cgtcccggt	ggttacgcta	tccctgaaatg	caacaacaaa	accttctccg	2580
gtaaagggtcc	gtgcaacaac	gtttccaccg	ttcagtgcac	ccatgggtatc	aaaccgggtg	2640
tttccaccca	gctgctgctg	aacgggtccc	tggctgaaaa	agaaatcatc	atccgttccg	2700
aaaacctgac	cgacaacggt	aaaaccatca	tcggttcacct	gaacaaatcc	gttgaaatcg	2760
tttgaccccc	tccgaacaac	aacacccgta	aatccatgcg	tatcgggtccg	ggtcagacct	2820
tctacgtac	cggtgacatc	atcggtgaca	tccgtcaggc	ttactgcaac	atctccgggt	2880
ccaaatggaa	cgaaaccctg	aaacgtgtta	aagaaaaact	gcaggaaaac	tacaacaaca	2940
acaaaacccat	caaattcgt	ccgtccctccg	gtgggtgacct	ggaaatcacc	accactcct	3000
tcaactgccg	tgggtgaattc	ttctactgca	acaccacccg	tctgttcaac	aacaacgcta	3060
ccgaagacga	aaccatcacc	ctgccgtgcc	gtatcaaaca	gatcatcaac	atgtggcagg	3120
gtgttggtcg	tgtatgtac	gctccgccga	tgcgtggtaa	catcacctgc	aaatccaaca	3180
tcaccgggtct	gctgctggtt	cgtgacgggtg	gtgaagacaa	caaaaccgaa	gaaatcttcc	3240
gtccgggtgg	tggtaacatg	aaagacaact	ggcgttccga	actgtacaaa	tacaaagtta	3300
tcgaactgaa	accgctgggt	atcgtccga	ccggtgctaa	gcttaccgtt	caggctcgtc	3360
agctgctgtc	ctccatcgtt	cagcagcagt	ccaacctgct	gcgtgctatc	gaagctcagc	3420
agcacatgct	gcagctgacc	gtttggggta	tcaaacagct	gcagaccctg	gttctggcta	3480
tcgaacgtta	cctgaaagac	cagcagctcg	agatctggaa	caacatgacc	tggatggaat	3540
gggaccgtga	aatctccaac	tacaccgaca	ccatctaccg	tctgctggaa	gactcccaga	3600
cccagcagga	aaaaaacgaa	aaagacctgc	tggctctgga	ctcctggaaa	aacctgtggt	3660
cctgggttcga	catctccaac	tggctgtggt	gaggatccag	atctgctgtg	ccttctagtt	3720
gccagccatc	tgttgtttgc	ccctcccccg	tgccttcctt	gacctggaa	gggtccactc	3780
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3840
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3900
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccaggt	gctgaagaat	tgacccggtt	3960
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	4020
cctgggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	4080
caatcccacc	cgctaaagta	cttgagagcg	tctctccctc	cctcatcagc	ccaccaaacc	4140
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	4200
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	4260
tgatttaagg	ccatcatggc	cttaatcttc	cgttccctcg	ctcactgact	cgtgcgctc	4320
ggtcgttcgg	ctgcggcgag	cggtatcagc	tactcaaaag	gcggtaatac	ggttatccac	4380
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4440
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgtccccctg	acgagcatca	4500
caaaaatcga	cgtcaagtc	agagggtggcg	aaacccgaca	ggactataaa	gataccaggc	4560

```

gtttccccct ggaagctccc tcgtgcgctc tctgtttccg accctgccgc ttaccggata 4620
cctgtccgcc ttttccccct cggaagcgt ggcgttttct catagctcac gctgtaggta 4680
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac ccccgttca 4740
gcccgaaccgc tgcgccttat ccggttaacta tcgtcttgag tccaaccccg taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 5040
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 5100
cgaaaactca cgttaaggga ttttggtcac gagattatca aaaaggatct tcacctagat 5160
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 5280
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggctc gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atttattcaa caaagccgcc gtcccgtaaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
tcaccgaggc agttccatag gatggcaaga tcttggtatc ggtctgcgat tccgactcgt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac aaccaaacg 5940
ttattcatc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatattt 6060
tcacctgaat caggatattc ttctaatacc tggaatgctg ttttccccgg gatcgcagtg 6120
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggctcg aagaggcata 6180
aattccgtca gccagttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgaatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6420
cttgattac tgtttatgta agcagacagt tttattgttc atgatgat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6718

```

&lt;210&gt; 128

&lt;211&gt; 6586

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 128

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggtggt cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420

```



```

ccccctggc tgaccgcccacgaccccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggggacgtcaatgg gtggagtatt tacggtaaacc 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcatatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccga ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatgggggc ccattttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtggggtc tccacgcgaa tctcgggtac gtgttccgga 1440
catggggtct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcgggtcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccgtgc tgttaacggg ggagggcaggt gtagtctgag cagtactcgt tgctgcccg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat ggggtcttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tctgtgtatc ccgcgttaact ggcgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgcgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca cctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2400
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggg aaaggctcgt 2460
gcaacaacgt tccaccgtt cagtgcaccc atggtatcaa accggttgtt tccaccagc 2520
tgctgctgaa cgggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcaccgctc 2640
cgaacaacaa caccgtaaa tccatgcgta tccgtccggg tcagaccttc tacgctaccg 2700
gtgacatcat cgggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2760
aaacctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaac aaaaccatca 2820
aattegtctc gtcctccggg ggtgacctgg aaatcaccac ccactccttc aactgccgtg 2880
gtgaattctt ctactgcaac accaccgctc tgttcaacaa caacgctacc gaagacgaaa 2940
ccatcaccct gccgtgccgt atcaaacaga tcatcaacat gtggcagggt gttggtcgtg 3000
ctatgtacgc tccgccgatc gctggtaaca tcacctgcaa atccaacatc accggtctgc 3060
tgctggttcg tgacgggtggg gaagacaaca aaaccgaaga aatcttccgt ccgggtgggtg 3120
gtaacatgaa agacaactgg cgttccgaac tgtacaaata caaagttatc gaactgaaac 3180
cgctgggtat cgctccgacc ggtgctaagc ttaccgttca ggctcgtcag ctgctgtcct 3240
ccatcgltca gcagcagtc aacctgctgc gtgctatcga agctcagcag cacatgctgc 3300
agctgaccgt ttgggggtatc aaacagctgc agaccggtg tctggctatc gaacgttacc 3360
tgaaagacca gcagctcgag atctggaaca acatgacctg gatggaatgg gaccgtgaaa 3420
tctccaacta caccgacacc atctaccgtc tgctggaaga ctcccagacc cagcaggaaa 3480
aaaacgaaaa agacctgctg gctctggact cctggaaaaa cctgtgggtcc tgggtcgaca 3540
tctccaactg gctgtggtga ggatccagat ctgctgtgcc ttctagttgc cagccatctg 3600
ttgtttgccc ctcccccggt ccttccttga ccctggaagg tgccactccc actgtccttt 3660
cctaataaaa tgaggaaatt gcacgcatt gtctgagtag gtgtcattct attctggggg 3720

```



gtggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	3780
atgcggtggg	ctctatgggt	accaggtgc	tgaagaattg	accggttcc	tcctgggcca	3840
gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	tccacgcccc	tggttcttag	3900
ttccagcccc	actcatagga	cactcatagc	tcaggagggc	tccgccttca	atcccacccg	3960
ctaaagtact	tggagcggtc	tctccctccc	tcatcagccc	accaaaccac	acctagcctc	4020
caagagtggg	aagaaattaa	agcaagatag	gctattaagt	gcagagggag	agaaaatgcc	4080
tccaacatgt	gaggaagtaa	tgagagaaat	catagaattt	taaggccatg	atttaaggcc	4140
atcatggcct	taatcttccg	cttctctgct	cactgactcg	ctgcgctcgg	tcgttcggct	4200
gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	4260
taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	4320
cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	4380
ctcaagtcag	aggtggcgaa	accgcagagg	actataaaga	taccaggcgt	ttccccctgg	4440
aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	4500
tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggtatc	tcagttcggg	4560
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	4620
cgccttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	4680
ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	4740
cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtattttggta	tctgcgctct	4800
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4860
cgctggtagc	ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4920
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	4980
ttaagggatt	ttgggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaat	5040
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttgggtctg	acagttacca	5100
atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	5160
ctgactcggg	gggggggggc	gctgaggtct	gcctcgtgaa	gaagggtgtg	ctgactcata	5220
ccaggcctga	atcgccccat	catccagcca	gaaagtgagg	gagccacggg	tgatgagagc	5280
tttgttgtag	gtggaccagt	tgggtgattt	gaacttttgc	tttgccacgg	aacggctctg	5340
gltgtcgga	agatgcgtga	tctgatcctt	caactcagca	aaagttcgat	ttattcaaca	5400
aagccgccgt	cccgtcaagt	cagcgtaatg	ctctgccagt	gttacaacca	attaaccaat	5460
tctgattaga	aaaactcatc	gagcatcaaa	tgaaactgca	atttattcat	atcaggatta	5520
tcaataccat	atttttgaaa	aagccgtttc	tgtaatgaag	gagaaaactc	accgaggcag	5580
ttccatagga	tggcaagatc	ctgggtatcg	tctgcgattc	cgactcgtcc	aacatcaata	5640
caacctatta	atttccctc	gtcaaaaata	aggttatcaa	gtgagaaatc	accatgagt	5700
acgactgaat	ccggtgagaa	tggcaaaagc	ttatgcattt	ctttccagac	ttgttcaaca	5760
ggccagccat	tacgctcgtc	atcaaaatca	ctcgcatcaa	ccaaaccgtt	attcattcgt	5820
gattgcgcct	gagcgagacg	aaatacgcg	tcgctgttaa	aaggacaatt	acaaacagga	5880
atcgaatgca	accggcgcag	gaacactgcc	agcgcatcaa	caatatcttc	acctgaatca	5940
ggatattctt	ctaatacctg	gaatgctgtt	ttcccgggga	tcgcagtggg	gagtaaccat	6000
gcatcatcag	gagtacggat	aaaatgcttg	atggtcggaa	gaggcataaa	ttccgtcagc	6060
cagtttagtc	tgaccatctc	atctgtaaca	tcattggcaa	cgctaccttt	gccatgtttc	6120
agaaacaact	ctggcgcatc	gggcttccca	tacaatcgat	agattgtcgc	acctgattgc	6180
ccgacattat	cgcgagccca	tttataccca	tataaatcag	catccatgtt	ggaatttaat	6240
cgcggcctcg	agcaagacgt	ttcccgttga	atatggctca	taacacccct	tgtattactg	6300
tttatgtaag	cagacagttt	tattgttcat	gatgatatat	ttttatcttg	tgcaatgtaa	6360
catcagagat	tttgagacac	aacgtggctt	ttccccccc	cccattattg	aagcatttat	6420
cagggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taaacaata	6480
ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	tctaagaaac	cattattatc	6540
atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	ttcgtc		6586

&lt;210&gt; 129

&lt;211&gt; 6505

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 129

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctcggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgcctcg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340
ccaccatcac ccaggcttgc ccgaaagtta acttcgacce gatcccgatc cactactgcg 2400
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggg aaaggtccgt 2460
gcaacaacgt ttccaccgtt cagtgcaccc atggatatcaa accggttgtt tccaccagc 2520
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcggt tgcacccggt 2640
ctagcgctta ctgcaacatc tccggttcca aatggaacga aacctgaaa cgtgttaaag 2700
aaaaactgca ggaaaactac aacaacaaca aaaccatcaa attcgctccg tcctccgggt 2760
gtgacctgga aatcaccacc cactccttca actgccgtgg tgaattcttc tactgcaaca 2820
ccaccgctct gttcaacaac aacgctaccg aagacgaaac catcacctg ccgtgccgta 2880
tcaaacagat catcaacatg tggcaggggt ttggctcgtg tatgtacgct ccgccgatcg 2940
ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctgggtcgt gacgggtggtg 3000

```

aagacaacaa	aaccgaagaa	atcttccgtc	cgggtggtgg	taacatgaaa	gacaactggc	3060
gttccgaact	gtacaaatac	aaagttatcg	aactgaaacc	gctgggtatc	gctccgaccg	3120
gtgctaagct	taccgttcag	gctcgtcagc	tgctgtcctc	catcgttcag	cagcagtcca	3180
acctgctgcg	tgctatcgaa	gctcagcagc	acatgctgca	gctgaccggt	tgggggatca	3240
aacagctgca	gacccgtggt	ctggctatcg	aacgttacct	gaaagaccag	cagctcgaga	3300
tctggaacaa	catgacctgg	atggaatggg	accgtgaaat	ctccaactac	accgacacca	3360
tctaccgtct	gctggaagac	tcccagaccc	agcaggaaaa	aaacgaaaaa	gacctgctgg	3420
ctctggactc	ctggaaaaac	ctgtggtcct	ggttcgacat	ctccaactgg	ctgtggtgag	3480
gatccagatc	tgctgtgcct	tctagtggcc	agccatctgt	tgtttgcccc	tcccccgctgc	3540
cttccttgac	cctggaaggt	gccactccca	ctgtcctttc	ctaataaaat	gaggaaattg	3600
catcgcatcg	tctgagtagg	tgctattcta	ttctgggggg	tgggggtggg	caggacagca	3660
agggggagga	ttgggaagac	aatagcaggc	atgctgggga	tgcggtgggc	tctatgggta	3720
cccagggtgct	gaagaattga	cccggttcct	cctggggccag	aaagaagcag	gcacatcccc	3780
ttctctgtga	cacaccctgt	ccacgcccc	ggttcttagt	tccagcccca	ctcataggac	3840
actcatagct	caggagggct	ccgccttcaa	tcccacccgc	taaagtactt	ggagcgggtct	3900
ctccctccct	catcagccca	ccaaacaaaa	cctagcctcc	aagagtggga	agaaattaaa	3960
gcaagatagg	ctattaagtg	cagagggaga	gaaaatgcct	ccaacatgtg	aggaagtaat	4020
gagagaaatc	atagaatttt	aaggccatga	tttaaggcca	tcatggcctt	aatcttccgc	4080
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	4140
ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	4200
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	4260
taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcgaaa	4320
cccagacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	4380
tgltccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	4440
gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcgctc	gctccaagct	4500
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	4560
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	4620
gattagcaga	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	4680
cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	4740
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	4800
tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	4860
ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggatcatgag	4920
attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	4980
ctaaagtata	tatgagtaaa	cttggctctga	cagttaccaa	tgcttaatca	gtgaggcacc	5040
tatctcagcg	atctgtctat	ttcggtcatc	catagttgcc	tgactcgggg	ggggggggcg	5100
ctgaggtctg	cctcgtgaag	aagggtgttg	tgactcatac	caggcctgaa	tgcgcccatc	5160
atccagccag	aaagtgaggg	agccacgggt	gatgagagct	ttgttgtagg	tggaccagtt	5220
gggtgatttt	aacttttgct	ttgccacgga	acggtctgcg	ttgtcgggaa	gatgcgtgat	5280
ctgatccctc	aactcagcaa	aagttcgatt	tattcaacaa	agccgccgct	ccgtcaagtc	5340
agcgtaatgc	tctgccagtg	ttacaaccaa	ttaaccaatt	ctgattagaa	aaactcatcg	5400
agcatcaaat	gaaactgcaa	tttatctata	tcaggattat	caataccata	tttttgaaaa	5460
agccgtttct	gtaatgaagg	agaaaactca	ccgaggcagt	tccataggat	ggcaagatcc	5520
tggtatcggt	ctgcgattcc	gactcgtcca	acatcaatac	aacctattaa	tttccccctc	5580
tcaaaaataa	ggttatcaag	tgagaaatca	ccatgagtga	cgactgaatc	cggtgagaat	5640
ggcaaaagct	tatgcatttc	tttccagact	tggtcaacag	gccagccatt	acgctcgtca	5700
tcaaaatcac	tgcgatcaac	caaaccgtta	ttcatctcgt	attgcgcctg	agcgagacga	5760
aatacgcgat	cgctgttaaa	aggacaatta	caaacaggaa	tcgaatgcaa	ccggcgcagg	5820
aacactgcca	gcgcgatcaac	aatattttca	cctgaatcag	gatattcttc	taatacctgg	5880
aatgctgttt	tcccggggat	cgcagtgggt	agtaaccatg	catcatcagg	agtaacggata	5940
aaatgcttga	tggtcggaag	aggcataaat	tccgtcagcc	agtttagtct	gaccatctca	6000
tctgtaacat	cattggcaac	gctacctttg	ccatgtttca	gaaacaactc	tggcgcatcg	6060
ggcttcccat	acaatcgata	gattgtcgca	cctgattgcc	cgacattatc	gcgagcccat	6120
ttatacccat	ataaatcagc	atccatgttg	gaatttaate	gcggcctcga	gcaagacggt	6180
tcccgttgaa	tatgggtcat	aacacccctt	gtattactgt	ttatgtaage	agacagtttt	6240
attgttcatg	atgatataat	tttatcttgt	gcaatgtaac	atcagagatt	ttgagacaca	6300

```

acgtggcctt ccccccccc ccattattga agcattttatc aggggttattg tctcatgagc 6360
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6420
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 6480
aggcgtatca cgaggccctt tcgtc                                     6505

```

&lt;210&gt; 130

&lt;211&gt; 6466

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 130

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcctg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtat ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacggtaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340

```



ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2400
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggg	aaagggtccg	2460
gcaacaacgt	ttccaccggt	cagtgcaccc	atggatatcaa	accgggttgt	tccaccacgc	2520
tgctgctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2580
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgaccccggt	2640
ctagcgctta	ctgcaacatc	tccgggtcca	aatggaacga	aacctgaaa	cgtgttaaag	2700
aaaaactgca	ggaaaactac	aacaacaaca	aaaccatcaa	attcgctccg	tcctccgggt	2760
gtgacctgga	aatcaccacc	cactccttca	actgccgtgg	tgaattcttc	tactgcaaca	2820
ccgctagcct	gccgtgccgt	atcaaacaga	tcacacacat	gtggcagggt	gttgggtcgt	2880
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	2940
tgctgggttcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggt	3000
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3060
cgctgggtat	cgctccgacc	gggtgctaag	tlaccgttca	ggctcgtcag	ctgctgtcct	3120
ccatcggtca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3180
agctgaccgt	ttgggggtatc	aaacagctgc	agaccctgtg	tctggctatc	gaacgttacc	3240
tgaaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3300
tctccaacta	caccgacacc	atctaccgtc	tgctggaaga	ctcccagacc	cagcaggaaa	3360
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtc	tggttcgaca	3420
tctccaactg	gctgtggtga	ggatccagat	ctgctgtgcc	ttctagtgtc	cagccatctg	3480
ttgtttgccc	ctcccccggt	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	3540
cctaataaaa	tgaggaaatt	gcacgcatt	gtctgagtag	gtgtcattct	attctggggg	3600
gtgggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	3660
atgcgggtggg	ctctatgggt	accaggtgc	tgaagaattg	accgggttcc	tcctggggcca	3720
gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	tcacgcccc	tggttcttag	3780
ttccagcccc	actcatagga	cactcatagc	tcaggagggc	tccgccttca	atcccacccg	3840
ctaaagtact	tggagcggtc	tctccctccc	tcacagccc	accaaacc	acctagcctc	3900
caagagtggg	aagaaattaa	agcaagatag	gctattaagt	gcagaggagg	agaaaatgcc	3960
tccaacatgt	gaggaagtaa	tgagagaaat	catagaattt	taaggccatg	atttaaggcc	4020
atcatggcct	taatcttccg	cttctctgct	cactgactcg	ctgcgctcgg	tcgttcggct	4080
gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	4140
taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	4200
cgcggttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	4260
ctcaagtcag	aggtggcgaa	accgcagagg	actataaaga	taccaggcgt	ttccccctgg	4320
aagctccctc	gtgcgctctc	ctgttccgac	cctgcccgtt	accggatacc	tgtccgcctt	4380
tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggatct	tcagttcggt	4440
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaacct	cccgttcagc	ccgaccgctg	4500
cgcttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	4560
ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	4620
cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtatttggta	tctgcgctct	4680
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4740
cgctggtagc	gggtggttttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4800
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaacg	aaaactcacg	4860
ttaagggtatt	ttgggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaat	4920
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttggctctg	acagttacca	4980
atgcttaate	agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	5040
ctgactcggg	gggggggggg	gctgagggtct	gcctcgtgaa	gaagggtgttg	ctgactcata	5100
ccaggcctga	atcgcccat	catccagcca	gaaagtgagg	gagccacggg	tgatgagagc	5160
tttggtgtag	gtggaccagt	tgggtgatttt	gaacttttgc	tttgccacgg	aacgggtctgc	5220
gttgctggga	agatgcgtga	tctgatcctt	caactcagca	aaagtctgat	ttattcaaca	5280
aagccgcctg	cccgtcaagt	cagcgtaaatg	ctctgccagt	gttacaacca	attaaccaat	5340
tctgattaga	aaaactcatc	gagcatcaaa	tgaaactgca	atttattcat	atcaggatta	5400
tcaataccat	atttttgaaa	aagccgtttc	tgtaatgaag	gagaaaactc	accgaggcag	5460
ttccatagga	tggcaagatc	ctggatcggg	tctgcgattc	cgactcgctc	aacatcaata	5520
caacctatta	atttccctc	gtcaaaaata	aggttatcaa	gtgagaaatc	accatgagtg	5580
acgactgaat	ccgggtgagaa	tggcaaaagc	ttatgcattt	ctttccagac	ttgttcaaca	5640



```

ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa ccaaaccgtt attcattcgt 5700
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5760
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 5820
ggatattctt ctaataacctg gaatgctgtt tccccgggga tcgcagtggg gagtaaccat 5880
gcatcatcag gaggacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 5940
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6000
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6060
ccgacattat cgcgagccca ttataacca tataaatcag catccatgtt ggaatttaat 6120
cgcggcctcg agcaagacgt ttcccggtga atatggctca taacaccctt tgtattactg 6180
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6240
catcagagat tttgagacac aacgtggctt tccccccccc cccattattg aagcatttat 6300
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6360
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6420
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6466

```

&lt;210&gt; 131

&lt;211&gt; 6547

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 131

```

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtcccgctt acataactta cggtaaattg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga cttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg acctatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgttt ggcaaaaaa tcaacgggac ttccaaaaa gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgccgt ccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtg cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

```

gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgtgt tgcgtggatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttcggtt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160
cgcaggaaat cgttctggaa aacgtttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2400
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaaggtccgt 2460
gcaacaacgt ttccaccggt cagtgcaccc atggtatcaa accggttgtt tccaccagc 2520
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcggt tgcacccgctc 2640
cgaacaacaa cacccgtaaa tccatgcgta tcgggtccggg tcagaccttc tacgctaccg 2700
gtgacatcat cggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2760
aaaccctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaac aaaaccatca 2820
aatctcgtcc gtcttccggt ggtgacctgg aaatcaccac ccactccttc aactgccgtg 2880
gtgaattctt ctactgcaac accgctagcc tgccgtgccg tatcaaacag atcatcaaca 2940
tgtggcaggg tgttggtcgt gctatgtacg ctccgccgat cgctggtaac atcacctgca 3000
aatccaacat caccggtctg ctgctgggtc gtgacgggtg tgaagacaac aaaaccgaag 3060
aaatcttccg tccgggtggg ggtaacatga aagacaactg gcgttccgaa ctgtacaaat 3120
acaaagttaa cgaactgaaa ccgctgggta tcgctccgac cgggtgctaag cttaccgttc 3180
aggctcgtca gctgctgtcc tccatcgctc agcagcagtc caacctgctg cgtgctatcg 3240
aagctcagca gcacatgctg cagctgaccg tttgggggat caaacagctg cagacccgtg 3300
ttctggctat cgaacgttac ctgaaagacc agcagctcga gatctggaac aacatgacct 3360
ggatggaatg ggaccgtgaa atctccaact acaccgacac catctaccgt ctgctggaag 3420
actcccagac ccagcaggaa aaaaacgaaa aagacctgct ggctctggac tcctggaaaa 3480
acctgtggtc ctgggttcgac atctccaact ggctgtgggt aggatccaga tctgctgtgc 3540
cttctagttg ccagccatct gttgtttgcc cctccccctg gccttccttg acctggaag 3600
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta 3660
gggtgtcatc tattctgggg ggtgggggtgg ggcaggacag caaggggggag gattgggaag 3720
acaatagcag gcatgctggg gatgcgggtg gctctatggg taccaggtg ctgaagaatt 3780
gacccgggtc ctctggggcc agaaagaagc aggcacatcc ccttctctgt gacacaccct 3840
gtccacgccc ctgggttctta gttccagccc cactcatagg acactcatag ctcaggaggg 3900
ctccgccttc aatcccaccc gctaaagtac ttggagcggg ctctccctcc ctcacagcc 3960
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4020
tgcagaggga gagaaaatgc ctccaacatg tgaggaaagta atgagagaaa tcatagaatt 4080
ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttccctcg tcactgactc 4140
gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct cactcaaagg cggtaatagc 4200
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag gccagcaaaa 4260
ggccaggaa cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4320
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4380
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4440
taccggatac ctgtccgcct ttctcccttc ggggaagcgt gcgctttctc atagctcacg 4500
ctgtaggtat ctgagttcgg tgtaggctcg tcgctccaag ctgggctgtg tgcacgaacc 4560
ccccgttcag cccgaccgct gcgccttate cggtaactat cgtcttgagt ccaacccggt 4620
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4680
tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4740
agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 4800
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 4860
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4920
tcagtggaa c gaaaactcac gtttaagggt tttggtcatg agattatcaa aaaggatctt 4980
cacctagatc ctttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5040

```

```

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5100
atbtcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc tgcctcgtga 5160
agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5220
ggagccacgg ttgatgagag ctttggttga ggtggaccag ttggtgattt tgaacttttg 5280
ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5340
aaaagttcga tttattcaac aaagccgccc tcccgtcaag tcagcgtaat gctctgccag 5400
tgttacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5460
aatattattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5520
ggagaaaact caccgaggca gttccatagg atggcaagat cctgggtatcg gtctgcgatt 5580
ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aagggttatca 5640
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaaag cttatgcatt 5700
tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcatca 5760
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5820
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5880
acaatatatt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 5940
atcgcagtggt tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6000
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6060
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6120
tagattgtcg cacctgattg cccgacatta tcgagagccc atttataccc atataaatca 6180
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6240
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6300
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tcccccccc 6360
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6420
atthagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6480
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6540
tttcgtc 6547

```

&lt;210&gt; 132

&lt;211&gt; 6637

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 132

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcgggggttg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc battaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggabtt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcccg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020

```

tccccgtgcc	aagagtgacg	taagtaccgc	ctatagactc	tataggcaca	ccccctttggc	1080
tcttatgcat	gctabactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccatttatta	tttacaaatt	cacatatata	acaacgcgct	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgtaaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tcgtgggtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgcttccgac	accaaagcct	2100
acgaccgtga	agttcacaac	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagccgtaac	tgctccttca	2340
acaccaccac	cgaaatccgt	gacaaaaaac	agcagggtta	cgctctgttc	taccgtccgg	2400
acatcgttct	gctgaaagaa	aaccgtaaca	actccaacaa	ctccgaatac	atcctgatca	2460
actgcaacgc	ttccaccatc	accagggctt	gcccgaaggt	taacttcgac	ccgatcccca	2520
tccactactg	cgctccggct	ggttacgcta	tcctgaaatg	caacaacaaa	accttctccg	2580
gtaaagggtcc	gtgcaacaac	gtttccaccg	ttcagtgcac	ccatgggtatc	aaaccggttg	2640
tttccacca	gctgctgctg	aacggttccc	tggctgaaaa	agaaatcatc	atccgttccg	2700
aaaacctgac	cgacaacggt	aaaaccatca	tcgttcacct	gaacaaatcc	gttgaaatcg	2760
tttgaccccg	tgctagcgct	tactgcaaca	tctccggttc	caaatggaac	gaaaccctga	2820
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	2880
cgctcctccg	tggtgacctg	gaaatcacca	cccactcctt	caactgccgt	ggatgaattct	2940
tctactgcaa	caccaccgct	ctgttcaaca	acaacgctac	cgaagacgaa	accatcacc	3000
tgccgtgccg	tatcaaacag	atcatcaaca	tgtggcaggg	tgttggctcg	gctatgtacg	3060
ctccgccgat	cgctggtaac	atcacctgca	aatccaacat	caccggctctg	ctgctggttc	3120
gtgacgggtg	tgaagacaac	aaaaccgaag	aaatcttccg	tccgggtggg	ggtaacatga	3180
aagacaactg	gcgttccgaa	ctgtacaaat	acaaagttat	cgaactgaaa	ccgctgggta	3240
tcgctccgac	cgggtgctaag	cttaccgttc	aggctcgtca	gctgctgtcc	tccatcgttc	3300
agcagcagtc	caacctgctg	cgtgctatcg	aagctcagca	gcacatgctg	cagctgaccg	3360
tttgggggtat	caaacagctg	cagaccgctg	ttctggctat	cgaacgttac	ctgaaagacc	3420
agcagctcga	gatctggaac	aacatgacct	ggatggaatg	ggaccgtgaa	atctccaact	3480
acaccgacac	catctaccgt	ctgctggaag	actcccagac	ccagcaggaa	aaaaacgaaa	3540
aagacctgct	ggctctggac	tcctggaaaa	acctgtggtc	ctggttcgac	atctccaact	3600
ggctgtgggtg	aggatccaga	tctgctgtgc	cttctagttg	ccagccatct	gttgtttgcc	3660
cctcccccg	gccttccttg	accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	3720
atgaggaaat	tgcatacgat	tgtctgagta	gggtgtcatc	tattctgggg	gggtggggtgg	3780
ggcaggacag	caaggggggag	gattgggaag	acaatagcag	gcatgctggg	gatgcgggtg	3840
gctctatggg	taccaggtg	ctgaagaatt	gacccgggtc	ctcctggggc	agaaagaagc	3900
aggcacatcc	ccttctctgt	gacacaccct	gtccacgccc	ctggttctta	gttccagccc	3960
cactcatagg	acactcatag	ctcaggaggg	ctccgccttc	aatccccacc	gctaaagtac	4020
ttggagcggg	ctctccctcc	ctcatcagcc	caccaaacca	aacctagcct	ccaagagtgg	4080
gaagaaatta	aagcaagata	ggctatatta	tgcagaggga	gagaaaatgc	ctccaacatg	4140
tgaggaagta	atgagagaaa	tcatagaatt	ttaaggccat	gatttaaggc	catcatggcc	4200
ttaatcttcc	gcttccctcg	tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	4260
ggtatcagct	cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	4320



```

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4380
ggcgtttttc cataggtccc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440
gaggtggcga aaccgcacag gactataaag ataccaggcg tttccccctg gaagctccct 4500
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4560
gggaagcgtg gcgctttctc atagctcacg ctgtagggtat ctgagttcgg tgtagggtcgt 4620
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttate 4680
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4740
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4800
gtggcctaac tacggctaca ctagaagaac agtatttgggt atctgcgctc tgctgaagcc 4860
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4980
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040
tttgggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5160
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5220
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5280
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttggttga 5340
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgtcggg 5400
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5460
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaacca ttctgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5700
aatttcccct cgtcaaaaat aagggttatca agtgagaaat caccatgagt gacgactgaa 5760
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820
ttacgctcgt catcaaaatc actcgcacat accaaaccgt tattcattcg tgattgcgcc 5880
tgagcgagac gaaatacgcg atcgtctgtta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcacat acaatatttt cacctgaatc aggatattct 6000
tctaatacct ggaatgctgt tttcccgggg atcgcagtggt tgagtaacca tgcacatca 6060
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6240
tcgcgagccc atttataccc atataaatca gcatccatgt tggaaattta tcgcggcctc 6300
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtgggt tccccccccc cccattatt gaagcattta tcagggttat 6480
tgtctcatga gcgatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600
acctataaaa ataggcgtat caccaggccc tttcgtc 6637

```

&lt;210&gt; 133

&lt;211&gt; 6598

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 133

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240

```



ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaattg	420
cccgccctggc	tgaccgcccc	acgacccccg	cccatlgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggccccgct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgcg	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccatttga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcgc	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggtc	ccatttatta	tttaciaaatt	cacatataca	acaacgccgt	ccccgtgcc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgtcctg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgccgtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcatgcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggatatcctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggg	ttggaccgac	gctaaaacca	ccctgttctg	cgttcccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagccgtaac	tgctccttca	2340
acaccaccac	cgaaatccgt	gacaaaaaac	agcagggtta	cgtctgttcc	taccgtccgg	2400
acatcgttct	gctgaaagaa	aaccgtaaca	actccaacaa	ctccgaatac	atcctgatca	2460
actgcaacgc	ttccaccatc	accaggctt	gcccgaaggt	taacttcgac	ccgatccccga	2520
tccactactg	cgtcccggt	ggttacgcta	tctgaaatg	caacaacaaa	accttctccg	2580
gtaaagggtcc	gtgcaacaac	gtttccaccg	ttcagtgcac	ccatgggtatc	aaaccgggtg	2640
tttccaccca	gctgctgctg	aacggttccc	tggctgaaaa	agaaatcatc	atccgttccg	2700
aaaacctgac	cgacaacgtt	aaaaccatca	tctgttcacct	gaacaaatcc	gttgaaatcg	2760
tttgcacccg	tgctagcgct	tactgcaaca	tctccgggtc	caaatggaac	gaaaccctga	2820
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	2880
cgtcctccgg	tgggtgacctg	gaaatcacca	cccactcctt	caactgccgt	gggtgaattct	2940
tctactgcaa	caccgctagc	ctgccgtgcc	gtatcaaaca	gatcatcaac	atgtggcagg	3000
gtgttgggtc	tgctatgtac	gctccgccga	tcgttggtaa	catcacctgc	aaatccaaca	3060
tcaccgggtct	gctgctgggt	cgtgacgggtg	gtgaagacaa	caaaaccgaa	gaaatcttcc	3120
gtccgggtgg	tggtaacatg	aaagacaact	ggcgttccga	actgtacaaa	tacaaagtta	3180
tcgaactgaa	accgctgggt	atcgttccga	ccggtgctaa	gcttaccgtt	caggctcgtc	3240
agctgctgtc	ctccatcggt	cagcagcagt	ccaacctgct	gcgtgctatc	gaagctcagc	3300
agcacatgct	gcagctgacc	gtttggggta	tcaaacagct	gcagaccctg	gttctggcta	3360
tcgaacgtta	cctgaaagac	cagcagctcg	agatctggaa	caacatgacc	tggatggaat	3420
gggaccgtga	aatctccaac	tacaccgaca	ccatctaccg	tctgctggaa	gactcccaga	3480
cccagcagga	aaaaaacgaa	aaagacctgc	tggctctgga	ctcctggaaa	aacctgtggt	3540

cctgggttcga catctccaac tggctgtggt gaggatccag atctgctgtg ccttctagtt 3600  
gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa ggtgccactc 3660  
ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 3720  
ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 3780  
ggcatgctgg ggatgcggtg ggctctatgg gtaccaggt gctgaagaat tgaccgggtt 3840  
cctcctgggc cagaaagaag caggcacatc cccttctctg tgacacaccc tgtccacgcc 3900  
cctggttcct agttccagcc ccactcatag gacactcata gctcaggagg gctccgcctt 3960  
caatcccacc cgctaaagta cttggagcgg tctctccctc cctcatcagc ccaccaaacc 4020  
aaacctagcc tccaagagt ggaagaaatt aaagcaagat aggctattaa gtgcagaggg 4080  
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat ttttaaggcca 4140  
tgatttaagg ccatcatggc cttaatcttc cgcttctctg ctactgact cgctgcgctc 4200  
ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaag gcggttaatac ggttatccac 4260  
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4320  
cggtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 4380  
caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc 4440  
gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 4500  
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4560  
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac ccccggttca 4620  
gcccgaaccg tgcgccttat ccggttaacta tcgtcttgag tccaacccgg taagacacga 4680  
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4740  
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4800  
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4860  
caaacaaacc accgctggta gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag 4920  
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 4980  
cgaaaactca cgttaaggga ttttggctcat gagattatca aaaaggatct tcacctagat 5040  
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5100  
tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc tatttctgtt 5160  
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5220  
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5280  
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5340  
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 5400  
atltattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5460  
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5520  
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5580  
tcaccgagge agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5640  
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc aagttagaaa 5700  
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5760  
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5820  
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5880  
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatat 5940  
tcacctgaat caggatatc ttctaatacc tggaatgctg ttttcccggg gatcgagtg 6000  
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggctcg aagaggcata 6060  
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6120  
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6180  
gcacctgatt gcccgaacat atcgcgagcc catttatacc catataaatc agcatccatg 6240  
ttggaattta atcgcgccct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6300  
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6360  
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccatat 6420  
tgaagcattt atcagggtta ttgtctcatg agcggataca tttttgaatg tatttagaaa 6480  
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6540  
accattatta tcatgacatt aacctataaa aataggcgta tcacgagggc ctttctgc 6598

&lt;211&gt; 6679

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 134

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc accgcccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc tggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatth gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttatth accattatth accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatatata acaacgccgt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccgtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaaacca cctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagccgtaac tgctccttca 2340
acaccaccac cgaaatccgt gacaaaaaac agcagggtta cgctctgttc taccgtccgg 2400
acatcgttct gctgaaagaa aaccgtaaca actccaacaa ctccgaatac atcctgatca 2460
actgcaacgc ttccaccatc acccaggett gccgaaagt taacttcgac ccgatcccg 2520
tccactactg cgctccggct ggttacgcta tcctgaaatg caacaacaaa accttctccg 2580
gtaaagggtc gtgcaacaac gtttccaccg ttcagtgcac ccatgggtatc aaaccggttg 2640
tttccacca gctgctgctg aacggttccc tggctgaaaa agaaatcatc atccgttccg 2700
aaaacctgac cgacaacgtt aaaacctca tegtacacct gaacaaatcc gttgaaatcg 2760
tttgcacccg tccgaacaac aacacccgta aatccatgcy tatcggtccg ggtcagacct 2820

```



tctacgctac	cggtgacatc	atcggtgaca	tccgtcagge	ttactgcaac	atctccggtt	2880
ccaaatggaa	cgaaccctg	aaacgtgtta	aagaaaaact	gcaggaaaac	tacaacaaca	2940
acaaaaccat	caaattcgct	ccgtcctccg	gtggtgacct	ggaaatcacc	accactcct	3000
tcaactgccg	tggatgaattc	ttctactgca	acaccgctag	cctgccgtgc	cgtatcaaac	3060
agatcatcaa	catgtggcag	ggtgttggtc	gtgctatgta	cgctccgccg	atcgctggta	3120
acatcacctg	caaattccaac	atcaccggtc	tgctgctggt	tcgtgacggt	ggtgaagaca	3180
acaaaaccga	agaaatcttc	cgctccgggtg	gtggtaacat	gaaagacaac	tggcggtccg	3240
aactgtacaa	atacaaagtt	atcgaactga	aaccgctggg	tatcgctccg	accggtgcta	3300
agcttaccgt	tcaggctcgt	cagctgctgt	cctccatcgt	tcagcagcag	tccaacctgc	3360
tgcgtgctat	cgaagctcag	cagcacatgc	tgcagctgac	cgtttggggt	atcaaacagc	3420
tgcagacccg	tgttctggct	atcgaacggt	acctgaaaga	ccagcagctc	gagatctgga	3480
acaacatgac	ctggatggaa	tgggaccgtg	aaatctccaa	ctacaccgac	accatctacc	3540
gtctgctgga	agactcccag	accagcagg	aaaaaaacga	aaaagacctg	ctggctctgg	3600
actcctggaa	aaacctgtgg	tcctgggttcg	acatctccaa	ctggctgtgg	tgaggatcca	3660
gatctgctgt	gccttctagt	tgccagccat	ctgttgtttg	cccctcccc	gtgccttcct	3720
tgacctgga	aggtgccact	cccactgtcc	tttccctaata	aaatgaggaa	attgcatcgc	3780
attgtctgag	taggtgtcat	tctattctgg	ggggtgggg	ggggcaggac	agcaaggggg	3840
aggattggga	agacaatagc	aggcatgctg	gggatgcggt	gggctctatg	ggtaccagg	3900
tgctgaagaa	ttgaccgggt	tcctcctggg	ccagaaagaa	gcaggcacat	ccccttctct	3960
gtgacacacc	ctgtccacgc	ccctgggttct	tagttccagc	cccactcata	ggacactcat	4020
agctcaggag	ggctccgcct	tcaatcccac	ccgctaaagt	acttgagcgc	gtctctccct	4080
ccctcatcag	cccaccaaac	caaacctagc	ctccaagagt	gggaagaaat	taaagcaaga	4140
taggctatta	agtgcagagg	gagagaaaat	gcctccaaca	tgtgaggaag	taatgagaga	4200
aatcatagaa	ttttaaggcc	atgatttaag	gccatcatgg	ccttaatctt	ccgcttctct	4260
gctcactgac	tcgctgcgct	cggtcggttcg	gctgcggcga	gcggtatcag	ctcactcaaa	4320
ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	tgtgagcaaa	4380
aggccagcaa	aaggccagga	accgtaaaaa	ggccgcggtg	ctggcggtttt	tccataggct	4440
ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	gaaacccgac	4500
aggactataa	agataccagg	cgtttcccc	tggaagctcc	ctcgtgcgct	ctcctgttcc	4560
gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	tggcgctttc	4620
tcatagctca	cgctgtaggt	atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	4680
tgtgcacgaa	cccccggttc	agcccagaccg	ctgcgcctta	tccggtaact	atcgtcttga	4740
gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	4800
cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tgggtggccta	actacggcta	4860
cactagaaga	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	tcggaaaaag	4920
agttggtagc	tcttgatccg	gcaaaacaaac	caccgctggt	agcgggtggtt	tttttggttg	4980
caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gatcctttga	tcttttctac	5040
ggggtctgac	gctcagtggg	acgaaaactc	acgttaaggg	attttggtca	tgagattatc	5100
aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	agttttaaat	caatctaaag	5160
tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	atcagtgagg	cacctatctc	5220
agcgatctgt	ctatttcggt	catccatagt	tgcctgactc	gggggggggg	ggcgctgagg	5280
tctgcctcgt	gaagaagggtg	ttgctgactc	ataccaggcc	tgaatcgccc	catcatccag	5340
ccagaaagtg	aggagaccac	ggttgatgag	agctttgttg	taggtggacc	agttggtgat	5400
tttgaacttt	tgctttgcca	cggaacggtc	tgcgttgtcg	ggaagatgcg	tgatctgac	5460
cttcaactca	gcaaaagttc	gattttattca	acaaagccgc	cgctccgctca	agtcagcgta	5520
atgctctgcc	agtgttacaa	ccaattaacc	aattctgatt	agaaaaactc	atcgagcatc	5580
aaatgaaact	gcaatttatt	catatcagga	ttatcaatac	catatttttg	aaaaagccgt	5640
ttctgtaatg	aaggagaaaa	ctcaccgagg	cagttccata	ggatggcaag	atcctggtat	5700
cggctctgcga	ttccgactcg	tccaacatca	atacaacctc	ttaattttcc	ctcgtcaaaa	5760
ataagggtat	caagtgagaa	atcaccatga	gtgacgactg	aatccgggtga	gaatggcaaa	5820
agcttatgca	tttctttcca	gacttggtca	acaggccagc	cattacgctc	gtcatcaaaa	5880
tcactcgcat	caaccaaac	gttattcatt	cgtgattgcg	cctgagcgag	acgaaatacg	5940
cgatcgctgt	taaaaggaca	attacaaaca	ggaatcgaat	gcaaccggcg	caggaacact	6000
gccagcgcat	caacaatatt	ttcacctgaa	tcaggatatt	cttctaatac	ctggaatgct	6060
gttttcccg	ggatcgcagt	ggtgagtaac	catgcatcat	caggagtacg	gataaaatgc	6120

```

ttgatggtcg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcactctgta 6180
acatcattgg caacgctacc ttbgccatgt ttcagaaaca actctggcgc atcgggcttc 6240
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6300
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccg 6360
tgaatatggc tcataacacc ccttgatatta ctgtttatgt aagcagacag ttttattgtt 6420
catgatgata tattttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6480
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 6540
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaaa 6600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6660
atcacgaggc cctttcgtc

```

&lt;210&gt; 135

&lt;211&gt; 6646

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 135

```

tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt taaggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtag atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggtctt tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgcgtccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatcgcgg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980

```



gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgcttccgac	accaaagcct	2100
acgaccgtga	agttcacaac	gtttgggcta	cccacgcttg	cgttccgacc	gaccggaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gctccttcgc	atgcatcaac	tgcaacgctt	2400
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2460
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaagggtccgt	2520
gcaacaacgt	ttccaccggt	cagtgcaccc	atggtatcaa	accggttggt	tccaccagc	2580
tgctgctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2640
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgaccccgtc	2700
cgaacaacaa	cacccgtaaa	tccatgcgta	tccggtccggg	tcagaccttc	tacgctaccg	2760
gtgacatcat	cggtgacatc	cgtcaggctt	actgcaacat	ctccggttcc	aaatggaacg	2820
aaaccctgaa	acgtgttaaa	gaaaaactgc	aggaaaacta	caacaacaac	aaaaccatca	2880
aattcgctcc	gtcctccggt	ggtgacctgg	aaatcaccac	ccactccttc	aactgccgtg	2940
gtgaattctt	ctactgcaac	accaccgctc	tgttcaacaa	caacgctacc	gaagacgaaa	3000
ccatcacccct	gccgtgccgt	atcaaacaga	tcatacaacat	gtggcagggt	gttgggtcgtg	3060
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	3120
tgctggttcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3180
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3240
cgctgggtat	cgctccgacc	ggtgctaagc	ttaccgttca	ggctcgtcag	ctgctgtcct	3300
ccatcgttca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3360
agctgaccgt	ttgggggtatc	aaacagctgc	agaccggtgt	tctggctatc	gaacgttacc	3420
tgaaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3480
tctccaacta	caccgacacc	atctaccgtc	tgctggaaga	ctcccagacc	cagcaggaaa	3540
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtcc	tggttcgaca	3600
tctccaactg	gctgtggtga	ggatccagat	ctgctgtgcc	ttctagtgtc	cagccatctg	3660
ttgtttgccc	ctcccccggtg	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	3720
cctaataaaa	tgaggaaatt	gcategcatt	gtctgagtag	gtgtcattct	attctggggg	3780
gtgggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	3840
atgcggtggg	ctctatgggt	accaggtgc	tgaagaattg	accgggttcc	tcctgggcca	3900
gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	tccacgcccc	tggttcttag	3960
ttccagcccc	actcatagga	cactcatagc	tcaggagggc	tccgccttca	atcccacccg	4020
ctaaagtact	tggagcggtc	tctccctccc	tcatacagccc	accaaaccaa	acctagcctc	4080
caagagtggg	aagaaattaa	agcaagatag	gctattaagt	gcagagggag	agaaaatgcc	4140
tccaacatgt	gaggaagtaa	tgagagaaat	catagaattt	taaggccatg	atttaaggcc	4200
atcatggcct	taatcttccg	cttccctcgt	cactgactcg	ctgcgctcgg	tcgttcgggt	4260
gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	4320
taacgcagga	aagaacatgt	gagcaaaaagg	ccagcaaaaag	gccaggaacc	gtaaaaaggc	4380
cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	4440
ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	4500
aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	4560
tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggatc	tcagttcgggt	4620
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	4680
cgccttatcc	ggtaactatc	gtcttgagtc	caacccggtg	agacacgact	tatcgccact	4740
ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	4800
cctgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtatttggtg	tctgcgctct	4860
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4920
cgctggtagc	ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4980
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	5040
ttaagggatt	ttgggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaatga	5100
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttgggtctg	acagttacca	5160
atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	tttcggtcat	ccatagttgc	5220
ctgactcggg	gggggggggg	gctgagggtc	gcctcgtgaa	gaagggtgtg	ctgactcata	5280

```

ccaggcctga atcgcccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5340
tttgttgtag gtggaccagt tgggtathtt gaacttttgc tttgccacgg aacggctctgc 5400
gttgtcggga agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca 5460
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5520
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5580
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5640
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5700
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5760
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5820
ggccagccat tacgctcgtc atcaaaatca ctgcgcatcaa ccaaaccggt attcattcgt 5880
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 5940
atcgaatgca accggcgcag gaacactgcc agcgcaccaa caatattttc acctgaatca 6000
ggatattctt ctaataacctg gaatgctgtt tccccgggga tcgcagtggg gagtaaccat 6060
gcatcatcag gagtacggat aaaatgcttg atggctcgga gaggcataaa ttccgtcagc 6120
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6180
agaaacaact ctggcgcacg gggcttccca tacaatcgat agattgtcgc acctgattgc 6240
ccgacattat cgcgagccca tttataccca tataaatcag catccatggt ggaatttaat 6300
cgccgcctcg agcaagacgt ttcccggtga atatggctca taacaccctt tgtattactg 6360
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6420
catcagagat tttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 6480
caggggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6540
gggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6600
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6646

```

&lt;210&gt; 136

&lt;211&gt; 6565

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 136

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgcctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttatbg accattattg accactcccc 1200

```

tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggtc	ccattttatta	tttacaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggatctctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtt	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacaac	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gctccttcgc	atgcatcaac	tgcaacgctt	2400
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2460
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggctcgt	2520
gcaacaacgt	ttccaccggt	cagtgcaccc	atggtatcaa	accggttggt	ttcaccacgc	2580
tgctgctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2640
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgaccccggt	2700
ctagcgctta	ctgcaacatc	tccggttcca	aatggaacga	aaccctgaaa	cgtgttaaag	2760
aaaaactgca	ggaaaactac	aacaacaaca	aaacctcaa	attcgctccg	tcctccggtg	2820
gtgacctgga	aatcaccacc	cactccttca	actgcctggg	tgaattcttc	tactgcaaca	2880
ccaccctct	gttcaacaac	aacgctaccg	aagacgaaac	catcaccctg	ccgtgccgta	2940
tcaaacagat	catcaacatg	tggcaggggt	ttggtcgtgc	tatgtacgct	ccgccgatcg	3000
ctggtaacat	cacctgcaaa	tccaacatca	ccggtctgct	gctgggtcgt	gacggtgggtg	3060
aagacaacaa	aaccgaagaa	atcttccgtc	cgggtgggtg	taacatgaaa	gacaactggc	3120
gttccgaact	gtacaaatac	aaagttatcg	aactgaaacc	gctgggtatc	gctccgaccg	3180
gtgctaagct	taccgttcag	gctcgtcagc	tgctgtcctc	catcgttcag	cagcagtcca	3240
acctgctgcg	tgctatcgaa	gctcagcagc	acatgctgca	gctgaccgtt	tggggtatca	3300
aacagctgca	gacccgtggt	ctggctatcg	aacgttacct	gaaagaccag	cagctcgaga	3360
tctggaacaa	catgacctgg	atggaatggg	accgtgaaat	ctccaactac	accgacacca	3420
tctaccgtct	gctggaagac	tcccagaccc	agcaggaaaa	aaacgaaaaa	gacctgctgg	3480
ctctggactc	ctggaaaaac	ctgtgggtcct	ggttcgacat	ctccaactgg	ctgtgggtgag	3540
gatccagatc	tgctgtgcct	tctagttgcc	agccatctgt	tgtttgcccc	ttccccgtgc	3600
cttccctgac	cctggaagggt	gccactccca	ctgtcctttc	ctaataaaaat	gaggaaattg	3660
catcgcattg	tctgagtagg	tgtcattcta	ttctgggggg	tgggggtggg	caggacagca	3720
agggggagga	ttgggaagac	aatagcaggc	atgctgggga	tgcgggtggg	tctatgggta	3780
cccagggtgct	gaagaattga	cccgggttcc	cctggggccag	aaagaagcag	gcacatcccc	3840
ttctctgtga	cacaccctgt	ccacgcccct	ggttcttagt	tccagcccca	ctcataggac	3900
actcatagct	caggaggggt	ccgccttcaa	tcccacccgc	taaagtactt	ggagcgggtct	3960
ctccctccct	catcagccca	ccaaaccaa	cctagcctcc	aagagtggga	agaaattaaa	4020
gcaagatagg	ctattaagtg	cagagggaga	gaaaatgcct	ccaacatgtg	aggaagtaat	4080
gagagaaatc	atagaatttt	aaggccatga	tttaaggcca	tcatggcctt	aatcttccgc	4140
ttcctcgtc	actgactcgc	tgcgtcgggt	cgttcggctg	cggcgagcgg	tatcagctca	4200
ctcaaaggcg	gtaatacgggt	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	4260
agcaaaaaggc	cagcaaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	4320
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	4380
cccgcagga	ctataaagat	accaggcgtt	tccccctgga	agctccctcg	tgcgctctcc	4440
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	4500

```

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 4560
gggctgtgtg cacgaacccc ccgttcagcc cgacgcgtgc gccttatccg gtaactatcg 4620
tcttgagtcc aaccgcggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 4680
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta 4740
cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4800
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 4860
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 4920
ttctacgggg tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgag 4980
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 5040
ctaaagtata tatgagtaaa cttggtctga cagttaccaaa tgcttaatca gtgaggcacc 5100
tatctcagcg atctgtctat ttctgttcac catagttgcc tgactcgggg gggggggggcg 5160
ctgaggtctg cctcgtgaag aagggtgtgc tgactcatat caggcctgaa tcgccccatc 5220
atccagccag aaagtgaggg agccacgggt gatgagagct ttgttgtagg tggaccagtt 5280
ggatgatttt aacttttgc tttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 5340
ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgcgctc ccgtcaagtc 5400
agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg 5460
agcatcaaat gaaactgcaa ttatttcata tcaggattat caataccata tttttgaaaa 5520
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 5580
tggtatcggg ctgcgattcc gactcgtcca acatcaatac aacctattaa tttccctcgg 5640
tcaaaaataa gggtatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 5700
ggcaaaaagc tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 5760
tcaaaatcac tcgcatcaac caaaccggtt ttcatctcgt attgcgcctg agcgagacga 5820
aatacgcgat cgctgtttaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 5880
aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 5940
aatgctgttt tcccggggat cgcagtggtg agtaaccatg catcatcagg agtacggata 6000
aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 6060
tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcacatc 6120
ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 6180
ttatacccat ataaatcagc atccatgttg gaatttaate gcggcctcga gcaagacggt 6240
tcccgttgaa tatgggtcat aacaccctt gtattactgt ttatgtaage agacagtttt 6300
attgttcatt atgatatatt tttatcttgt gcaatgtaac atcagagatt ttgagacaca 6360
acgtggcttt cccccccccc ccattattga agcatttatc agggttattg tctcatgagc 6420
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 6480
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaaat 6540
aggcgtatca cgaggccctt tcgtc

```

&lt;210&gt; 137

&lt;211&gt; 6526

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 137

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttgggtcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgctt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480

```



catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatbga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgcttg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcag	taagtaccgc	ctatagactc	tataggcaca	cccccttgge	1080
tcttatgcat	gctatactgt	ttttggcttg	gggctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatat	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggagc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcgagagc	tccacatccg	agccctgggc	ccatgcctcc	1500
agcggctcat	ggtcgcctcg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgtaaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgacgtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgtaact	ggcgcgagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttcccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gctccttcgc	atgcatcaac	tgcaacgctt	2400
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2460
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggctccg	2520
gcaacaacgt	ttccaccgtt	cagtgcaccc	atggatatcaa	accggttgtt	tccaccagc	2580
tgctgctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2640
acaacgttaa	aacctatcat	gttcacctga	acaaatccgt	tgaaatcggt	tgcacccgtg	2700
ctagcgctta	ctgcaacatc	tccggttcca	aatggaacga	aacctgaaa	cgtgttaaa	2760
aaaaactgca	ggaaaactac	aacaacaaca	aaacctcaaa	atcgcctccg	tcctccggtg	2820
gtgacctgga	aatcaccacc	cactccttca	actgccgtgg	tgaattcttc	tactgcaaca	2880
ccgctagcct	gccgtgccgt	atcaaacaga	tcacacacat	gtggcagggt	gttggctcgtg	2940
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	3000
tgctgggttcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3060
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3120
cgtgggtat	cgtccgacc	ggtgctaagc	ttaccttca	ggctcgtcag	ctgctgtcct	3180
ccatcgctta	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3240
agctgaccgt	ttggggatc	aaacagctgc	agaccggtgt	tctggctatc	gaacgttacc	3300
tgaaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3360
tctccaacta	caccgacacc	atctaccgtc	tgctggaaga	ctcccagacc	cagcaggaaa	3420
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtc	tggttcgaca	3480
tctccaactg	gctgtgggtg	ggatccagat	ctgctgtgcc	ttctagtgtc	cagccatctg	3540
ttgtttgccc	ctcccccggtg	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	3600
cctaataaaa	tgaggaaatt	gcacgcgatt	gtctgagtag	gtgtcattct	attctggggg	3660
gtgggggtggg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	3720
atgcgggtggg	ctctatgggt	accaggtgc	tgaagaattg	acccgggttc	tcctgggcca	3780



```

gaaagaagca ggcacatccc cttctctgtg acacaccctg tccacgcccc tggttcttag 3840
ttccagcccc actcatagga cactcatagc tcaggagggc tccgccttca atcccacccg 3900
ctaaagtact tggagcgggc tctccctccc tcatcagccc accaaaccaa acctagcctc 3960
caagagtggg aagaaattaa agcaagatag gctattaagt gcagagggag agaaaatgcc 4020
tccaacatgt gaggaagtaa tgagagaaat catagaattt taaggccatg atttaaggcc 4080
atcatggcct taatcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 4140
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 4200
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 4260
cgcgltgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 4320
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 4380
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 4440
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggg 4500
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 4560
cgcttctatc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact 4620
ggcagcagcc actggtaaca ggattagcag agcgaggatg tagggcgggtg ctacagagtt 4680
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct 4740
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 4800
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 4860
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 4920
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 4980
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggctcg acagttacca 5040
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 5100
ctgactcggg gggggggggc gctgaggtct gcctcgtgaa gaagggtgtg ctgactcata 5160
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc 5220
tttgtttagt gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc 5280
gttgtcggga agatgcgtga tctgacctt caactcagca aaagttcgat ttattcaaca 5340
aagccgccgt cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat 5400
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 5460
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 5520
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 5580
caacctatta atttccccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 5640
acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 5700
ggccagccat tacgctcgtc atcaaaatca ctgcgcatcaa ccaaaccgtt attcattcgt 5760
gattgcgcct gagcgagacg aaatacgcga tgcgtgttaa aaggacaatt acaaacagga 5820
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 5880
ggatattctt ctaataacctg gaatgctgtt ttcccgggga tgcgagtggg gagtaaccat 5940
gcatcatcag gactacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 6000
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 6060
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 6120
ccgacattat cgcgagccca tttataacca tataaatcag catccatgtt ggaatttaat 6180
cgcggcctcg agcaagacgt ttcccgttga atatggctca taacaccctt tgtattactg 6240
tttatgtaag cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa 6300
catcagagat tttgagacac aacgtggctt tccccccccc cccattattg aagcatttat 6360
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 6420
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 6480
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 6526

```

&lt;210&gt; 138

&lt;211&gt; 6607

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 138

```

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60
cagcttgtctt gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaate cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttaaaatt cacaataaca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgcgtccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttcgggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaa gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gtccttcgc atgcatcaac tgcaacgctt 2400
ccaccatcac ccaggcttgc ccgaaagtta acttcgacce gatcccgatc cactactgcg 2460
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac ctctcctcgg aaaggtccgt 2520
gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccaccagc 2580
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcggt tgcaccgctc 2700
cgaacaacaa caccgtaaa tccatgcgta tcgggtccgg tcagaccttc tacgtaccg 2760
gtgacatcat cggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2820
aaaccctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaa aaaaccatca 2880
aatcgcctcc gtcctccggt ggtgacctgg aaatcaccac ccactccttc aactgccgtg 2940
gtgaattctt ctactgcaac accgctagcc tgccgtgccg tatcaaacag atcatcaaca 3000
tgtggcaggg tgttggtcgt gctatgtacg ctccgccgat cgctggtaac atcacctgca 3060
aatccaacat caccggtctg ctgctggttc gtgacggtgg tgaagacaac aaaaccgaag 3120

```

```

aaatcttccg tccgggtggt ggtaacatga aagacaactg gcgttccgaa ctgtacaaat 3180
acaaagtatt cgaactgaaa ccgctgggta tcgctccgac cggtgctaag cttaccgttc 3240
aggctcgtca gctgctgtcc tccatcgttc agcagcagtc caacctgctg cgtgctatcg 3300
aagctcagca gcacatgctg cagctgaccg tttgggggat caaacagctg cagacccgtg 3360
ttctggctat cgaacgttac ctgaaagacc agcagctcga gatctggaac aacatgacct 3420
ggatggaatg ggaccgtgaa atctccaact acaccgacac catctaccgt ctgctggaag 3480
actcccagac ccagcaggaa aaaaacgaaa aagacctgct ggctctggac tcctggaaaa 3540
acctgtggtc ctggttcgac atctccaact ggctgtgggt aggatccaga tctgctgtgc 3600
cttctagttg ccagccatct gttgtttgcc cctccccgt gccttccttg accctggaag 3660
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcctcgcat tgtctgagta 3720
ggtgtcatte tattctgggg ggtgggggtgg ggcaggacag caaggggggag gattgggaag 3780
acaatagcag gcatgctggg gatgcggtgg gctctatggg taccaggtg ctgaagaatt 3840
gacccgggtc ctccctgggc agaaagaagc aggcacatcc ccttctctgt gacacacct 3900
gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag ctgaggagg 3960
ctccgccttc aatcccaccc gctaaagtac ttggagcggc ctctccctcc ctcatcagcc 4020
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4080
tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa tcatagaatt 4140
ttaaggccat gatttaagge catcatggcc ttaatcttcc gcttctctgc tcatgactc 4200
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4260
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4320
ggccagggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggtccc gccccctga 4380
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag 4440
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4500
taccggatac ctgtccgcct ttctcccttc ggggaagcgtg gcgctttctc atagctcacg 4560
ctgtaggtat ctgagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4620
ccccgttcag ccgaccgct gcgccttate cggtaactat cgtcttgagt ccaaccgggt 4680
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4740
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4800
agtatttggg atctgcgctc tgcgaagcc agttaccttc ggaaaaagag ttggtagctc 4860
ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat 4920
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4980
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5040
cacctagatc cttttaaatt aaaaatgaag ttttaaatac atctaaagta tataatgagta 5100
aacttgggtc gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5160
atttcggttc tccatagttg cctgactcgg gggggggggg cgctgagggtc tgcctcgtga 5220
agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagttag 5280
ggagccacgg ttgatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg 5340
ctttgccacg gaacgggtctg cgttgctcggg aagatgcgtg atctgatcct tcaactcagc 5400
aaaagttcga tttattcaac aaagccgccg tcccgtaag tcagcgtaat gctctgccag 5460
tggtacaacc aattaaccaa ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5520
aatttattca tatcaggatt atcaatacca tttttttgaa aaagccgttt ctgtaatgaa 5580
ggagaaaact caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt 5640
ccgactcgtc caacatcaat acaacctatt aatttccct cgtcaaaaat aaggttatca 5700
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5760
tctttccaga cttgttcaac aggcagcca ttacgctcgt catcaaaatc actcgcatca 5820
accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5880
aaaggacaat tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5940
acaatathtt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 6000
atcgcagtggt tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6060
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6120
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6180
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6240
gcatccatgt tggaatttaa tcgcggcctc gagcaagacg tttcccgttg aatatggctc 6300
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6360
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct ttccccccc 6420

```

```

ccccattatt gaagcatttta tcaggggttat tgtctcatga gcggatacat atttgaatgt 6480
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6540
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 6600
tttcgtc                                         6607

```

&lt;210&gt; 139

&lt;211&gt; 6676

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 139

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgct agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgccgtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtat ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaa gtttgggcta cccacgcttg cgttccgacc gaccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340

```



```

ccaacgacat gaacaaagaa atccgtaact gtccttcaa caccaccacc gaaatccgtg 2400
acaaaaaaca gcaggggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460
accgtaacaa ctccaacaac tccgaatata tcctgatcaa ctgcaacgct tccaccatca 2520
cccaggttg cccgaaagt aacttcgacc cgatcccgat ccactactgc gctccggctg 2580
gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaagggtccg tgcaacaacg 2640
tttccaccgt tcagtgcacc catggtatca aaccggttgt tttccaccag ctgctgctga 2700
acggttcctt ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760
aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcaccctg gctagcgctt 2820
actgcaacat ctccggttcc aaatggaacg aaaccttgaa acgtgtttaa gaaaaactgc 2880
aggaaaacta caacaacaac aaaaccatca aattcgctcc gtctccgggt ggtgacctgg 2940
aaatcaccac ccactccttc aactgccgtg gtgaattctt ctactgcaac accaccctgc 3000
tgttcaacaa caacgtacc gaagacgaaa ccataccctt gccgtgccgt atcaaacaga 3060
tcatcaacat gtggcagggg gttggtcgtg ctatgtacgc tccgccgatc gctggtaaca 3120
tcacctgcaa atccaacatc accggtctgc tgctgggttcg tgacgggtggg gaagacaaca 3180
aaaccgaaga aatcttccgt cgggtgggtg gtaacatgaa agacaactgg cgttccgaac 3240
tgtacaaata caaagttatc gaactgaaac cgctgggtat cgctccgacc ggtgctaagc 3300
ttaccgttca ggctcgtcag ctgctgtcct ccatacgttca gcagcagtec aacctgctgc 3360
gtgctatcga agctcagcag cacatgctgc agctgaccgt ttgggggtatc aaacagctgc 3420
agaccctgtg tctggctatc gaacgttacc tgaaagacca gcagctcgag atctggaaca 3480
acatgacctg gatggaatgg gaccgtgaaa tctccaacta caccgacacc atctaccgtc 3540
tgctggaaga ctcccagacc cagcaggaaa aaaacgaaaa agacctgctg gctctggact 3600
cctggaaaaa cctgtggtcc tgggttcgaca tctccaactg gctgtggtga ggatccagat 3660
ctgctgtgcc ttctagttgc cagccatctg ttgtttgcc ctccccctg ccttccttga 3720
ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatacgcat 3780
gtctgagtag gtgtcattct attctggggg gtgggggtggg gcaggacagc aagggggagg 3840
attgggaaga caatagcagg catgctgggg atgcgggtgg ctctatgggt acccaggtgc 3900
tgaagaattg acccgttcc tctggggcca gaaagaagca ggcacatccc ctctctctgt 3960
acacaccctg tccacgcccc tgggttcttag ttccagcccc actcatagga cactcatagc 4020
tcaggagggg tccgccttca atcccacccg cttaaagtact tggagcggtc tctccctccc 4080
tcatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4140
gctattaagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat 4200
catagaattt taaggccatg atttaaggcc atcatggcct taatcttccg ctctctcgt 4260
cactgactcg ctgcgctcgg tcgttcggct ggcgcgagcg gtatcagctc actcaaaggc 4320
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 4380
ccagcaaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 4440
ccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 4500
actataaaga taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac 4560
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 4620
tagctcacgc ttaggtatc tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt 4680
gcacgaaccc ccggttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 4740
caaccctgta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 4800
agcgagggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 4860
tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 4920
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgtttgcaa 4980
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 5040
gtctgacgct cagtggaacg aaaactcacg ttaagggtat ttggtcatga gattatcaaa 5100
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 5160
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 5220
gatctgtcta ttctgttcat ccatagttgc ctgactcggg gggggggggc gctgaggtct 5280
gcctcgtgaa gaagggtgtg ctgactcata ccaggcctga atcgccccat catccagcca 5340
gaaagtgagg gagccacggt tgatgagagc tttgttgtag gtggaccagt tgggtathtt 5400
gaacttttgc tttgccacgg aacggtctgc gttgtcggga agatgcgtga tctgatcctt 5460
caactcagca aaagtctgat ttattcaaca aagccgcgt cccgtcaagt cagcgtaatg 5520
ctctgccagt gttacaacca attaaccaat tctgattaga aaaactcacc gagcatcaaa 5580
tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc 5640

```



```

tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg 5700
tctgcgattc cgactcgtec aacatcaata caacctatta atttcccctc gtcaaaaata 5760
aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc 5820
ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca 5880
ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga 5940
tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag gaacactgcc 6000
agcgcatcaa caataatttc acctgaatca ggatattcct ctaataacctg gaatgctgtt 6060
ttcccgggga tcgcagtggg gagtaaccat gcatcatcag gagtacggat aaaatgcttg 6120
atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca 6180
tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcacg gggcttccca 6240
tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca ttataacca 6300
tataaatcag catccatgtt ggaatttaac cgcggcctcg agcaagacgt ttcccgttga 6360
atatggctca taacaccctt tgtattactg tttatgtaag cagacagttt tattgttcat 6420
gatgatatat ttttatcttg tgcaatgtaa catcagagat tttgagacac aacgtggctt 6480
tccccccccc ccattatttg aagcatttat cagggttatt gtctcatgag cggatacata 6540
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6600
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6660
acgaggccct ttcgtc 6676

```

&lt;210&gt; 140

&lt;211&gt; 6637

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 140

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500

```

agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctaga gatatcgcca 1920  
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980  
gtttctggat gatcatcatc tgccgtggtt ttggtaacat gtgggttacc gtttactacg 2040  
gtgttcgggt ttggaccgac gctaaaacca ccctgttctg cgcttcgcac accaaagcct 2100  
acgaccgtga agttcacac gtttgggcta cccacgcttg cgttcgcacc gacccgaacc 2160  
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220  
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280  
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340  
ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccaccacc gaaatccgtg 2400  
acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460  
accgtaacaa ctccaacaac tccgaataca tcctgatcaa ctgcaacgct tccaccatca 2520  
cccaggcttg cccgaaagtt aacttcgacc cgatcccgat ccactactgc gctccggctg 2580  
gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaaggctcg tgcaacaacg 2640  
tttccaccgt tcagtgacac catggtatca aaccggttgt ttccaccag ctgctgctga 2700  
acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760  
aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcacctgt gctagcgtt 2820  
actgcaacat ctccggttcc aaatggaacg aaacctgaa acgtgttaaa gaaaaactgc 2880  
aggaaaacta caacaacaac aaaaccatca aattcgctcc gtctccgggt ggtgacctgg 2940  
aaatcaccac ccactccttc aactgccgtg gtgaattctt ctactgcaac accgctagcc 3000  
tgccgtgccg tatcaaacag atcatcaaca tgtggcaggg tgttggtcgt gctatgtacg 3060  
ctccgccgat cgttggtaac atcacctgca aatccaacat caccggtctg ctgctggttc 3120  
gtgacggtgg tgaagacaac aaaaccgaag aaatcttccg tccgggtggg ggtaacatga 3180  
aagacaactg gcgttccgaa ctgtacaaat acaaagttat cgaactgaaa ccgctgggta 3240  
tcgctccgac cgggtgctaag cttaccgttc aggtcgtca gctgctgtcc tccatcgttc 3300  
agcagcagtc caacctgctg cgtgctatcg aagctcagca gcacatgctg cagctgaccg 3360  
tttgggggat caaacagctg cagaccctgt ttctggctat cgaacgttac ctgaaagacc 3420  
agcagctcga gatctggaac aacatgacct ggatggaatg ggaccgtgaa atctccaact 3480  
acaccgacac catctaccgt ctgctggaag actcccagac ccagcaggaa aaaaacgaaa 3540  
aagacctgct ggctctggac tcctggaaaa acctgtggtc ctggttcgac atctccaact 3600  
ggctgtggtg aggatccaga tctgctgtgc cttctagtgt ccagccatct gttgtttgcc 3660  
cctccccctg gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa 3720  
atgaggaaat tgcatcgcat tgtctgagta ggtgtcatte tattctgggg ggtgggggtg 3780  
ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg gatgcgggtg 3840  
gctctatggg taccaggtg ctgaagaatt gaccgggttc ctctggggcc agaaagaagc 3900  
aggcacatcc ccttctctgt gacacacct gtccacgccc ctggttctta gttccagccc 3960  
cactcatagg aactcatag ctgaggagg ctccgccttc aatcccaccc gctaaagtac 4020  
ttggagcggg ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg 4080  
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4140  
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4200  
ttaatcttcc gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4260  
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4320  
aaagaacatg tgagcaaaa ggcagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct 4380  
ggcgttttcc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440  
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500  
cgtgcgctct cctgttccga ccctgcccgt taccggatac ctgtccgcct ttctcccttc 4560  
gggaagcgtg gcgcttcttc atagctcacg ctgtaggtat ctgagttcgg tgtaggtcgt 4620  
tcgctccaag ctgggctgtg tgcacgaacc cccggttcag cccgaccgct gcgccttacc 4680  
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4740  
cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg 4800

```

gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 4860
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920
cgggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4980
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040
tttgggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5160
cagtgaaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5220
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5280
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 5340
gggtggaccag ttgggtgattt tgaacttttg ctttgccacg gaacgggtctg cgttgtcggg 5400
aagatgcgtg atctgacctt tcaactcagc aaaagtctga tttattcaac aaagccgccg 5460
tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5700
aatttccctt cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 5760
tccgggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820
ttacgctcgt catcaaaatc actcgcacat accaaaccgt tattcattcg tgattgcgcc 5880
tgagcgagac gaaatacgcg atcgtgttta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcacat acaatatttt cacctgaatc aggatattct 6000
tctaatacct ggaatgctgt tttcccgggg atcgcagttg tgagtaacca tgcacatca 6060
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tctggcgcac cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6240
tcgcgagccc atttataccc atataaatca gcatccatgt tggaatttaa tcgcggcctc 6300
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtggct tcccccccc cccattatt gaagcattta tcagggttat 6480
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600
acctataaaa ataggcgtat caccaggccc tttcgtc 6637

```

&lt;210&gt; 141

&lt;211&gt; 6718

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 141

```

tcgcgcggttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtgaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtcccgctt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720

```

catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagccctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgaag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcgaggct	tccacatccg	agccctggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtggtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgtaaacggt	ggaggcgagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacac	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgttaact	gctccttcaa	caccaccacc	gaaatccgtg	2400
acaaaaaaca	gcagggttac	gctctgttct	accgtccgga	catcgttctg	ctgaaagaaa	2460
accgtaacaa	ctccaacaac	tccgaataca	tcctgatcaa	ctgcaacgct	tccaccatca	2520
cccaggcttg	cccgaaggtt	aacttcgacc	cgatcccgat	ccactactgc	gctccggttg	2580
gttacgctat	cctgaaatgc	aacaacaaaa	ccttctccgg	taaaggctcc	tgcaacaacg	2640
tttccaccgt	tcagtgcacc	catggtatca	aaccggttgt	ttccaccag	ctgctgctga	2700
acggttccct	ggctgaaaaa	gaaatcatca	tccgttccga	aaacctgacc	gacaacgtta	2760
aaaccatcat	cgttcacctg	aacaaatccg	ttgaaatcgt	ttgcacctgt	ccgaacaaca	2820
acaccgtaaa	atccatgcgt	atcgggtccg	gtcagacctt	ctacgctacc	ggtgacatca	2880
tccgtgacat	ccgtcaggct	tactgcaaca	tctccggttc	caaattggaac	gaaaccttga	2940
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	3000
cgtcctccgg	tggtagacctg	gaaatcacca	cccactcctt	caactgccgt	ggtgaattct	3060
tctactgcaa	caccgctagc	ctgccgtgcc	gtatcaaaca	gatcatcaac	atgtggcagg	3120
gtgttggtcg	tgctatgtac	gctccgccga	tccgtgtgta	catcacctgc	aaatccaaca	3180
tcaccggtct	gctgctgggt	cgtgacgggtg	gtgaagacaa	caaaaccgaa	gaaatcttcc	3240
gtccgggtgg	tggtaacatg	aaagacaact	ggcgttccga	actgtacaaa	tacaaagtta	3300
tcgaactgaa	accgctgggt	atcgtcccca	ccggtgctaa	gcttaccggt	caggctcgtc	3360
agctgctgtc	ctccatcggt	cagcagcagt	ccaacctgct	gcgtgctatc	gaagctcagc	3420
agcacatgct	gcagctgacc	gtttggggta	tcaaacagct	gcagacctgt	gttctggcta	3480
tcgaacgtta	cctgaaagac	cagcagctcg	agatctggaa	caacatgacc	tggatggaat	3540
gggaccgtga	aatctccaac	tacaccgaca	ccatctaccg	tctgctggaa	gactcccaga	3600
cccagcagga	aaaaaacgaa	aaagacctgc	tggctctgga	ctcctggaaa	aacctgtggt	3660
cctgggttcga	catctccaac	tggctgtggt	gaggatccag	atctgctgtg	ccttctagtt	3720
gccagccatc	tggtgtttgc	ccctcccccg	tgccttccct	gacctgggaa	ggtgccactc	3780
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3840
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3900
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccaggtg	gctgaagaat	tgacctgggt	3960
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgctccacgc	4020



```

cctgggttctt agttccagcc ccactcatag gacactcata gctcaggagg gctccgcctt 4080
caatcccacc cgctaaagta cttggagcgg tctctccctc cctcatcagc ccaccaaac 4140
aaacctagcc tccaagagtg ggaagaaatt aaagcaagat aggctattaa gtgcagaggg 4200
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat tttaaggcca 4260
tgattttaagg ccatcatggc cttaatcttc cgcttccctc ctcactgact cgctgcgctc 4320
ggtcggttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatat gggtatccac 4380
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4440
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 4500
caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc 4560
gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 4620
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4680
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac ccccgttca 4740
gcccagaccgc tgcgccttat ccggttaacta tcgtcttgag tccaacccgg taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag 5040
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 5100
cgaaaactca cgttaaggga ttttgggtcat gagattatca aaaaggatct tcacctagat 5160
cctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc tatttcgttc 5280
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatte 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
tcaccgaggg agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac aaccaaaccc 5940
ttattcatte gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatattt 6060
tcacctgaat caggatatte ttctaatacc tggaatgctg ttttcccggg gatcgcagtg 6120
gtgagtaacc atgcatcacc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6180
aattccgtca gccagtttag tctgaccatc tcactctgtaa catcattggc aacgctacct 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcgccct cgagcaagac gtttcccggt gaatatggct cataacaccc 6420
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccattat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6718

```

&lt;210&gt; 142

&lt;211&gt; 6790

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes



&lt;400&gt; 142

```

tcgcgcggttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcgtt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatgcctg gagacgccat ccacgtgtt ttgacctca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggctt gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggagc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctcctt ctcctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggaggcagc gtatgtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagccgtaac tgctccttca 2340
acaccaccac cgaaatccgt gacaaaaaac agcagggtta cgctctgttc taccgtccgg 2400
acatcgctct gctgaaagaa aaccgtaaca actccaacaa ctccgaatac atcctgatca 2460
actgcaacgc tccaccatc acccaggctt gcccgaaagt taacttcgac ccgatcccga 2520
tccactactg cgctccggct ggttacgcta tcctgaaatg caacaacaaa accttctccg 2580
gtaaagggtc gtgcaaacac gtttccaccg ttcagtgcac ccatggatc aaaccggttg 2640
tttccacca gctgctgctg aacggttccc tggctgaaaa agaaatcatc atccgttccg 2700
aaaacctgac cgacaacgtt aaaaccatca tcgttcacct gaacaaatcc gttgaaatcg 2760
tttgaccccg tccgaacaac aacaccgta aatccatgcg tatcggtcgg ggtcagacct 2820
tctacgctac cggtgacatc atcggtgaca tccgtcaggc ttactgcaac atctccggtt 2880
ccaaatggaa cgaaaccctg aaacgtgtta aagaaaaact gcaggaaaac tacaacaaca 2940
acaaaaccat caaatcgtt ccgtcctccg gtggtgacct ggaaatcacc acccactcct 3000
tcaactgccg tgggtgaattc ttctactgca acaccaccg tctgttcaac aacaacgcta 3060
ccgaagacga aaccatcacc ctgccgtgcc gtatcaaaca gatcatcaac atgtggcagg 3120
gtgttggtcg tgctatgtac gctccgccga tcgctggtaa catcacctgc aaatccaaca 3180

```

```

tcaccggtct gctgctgggt cgtgacgggt gtgaagacaa caaaaccgaa gaaatcttcc 3240
gtccgggtgg tggtaacatg aaagacaact ggcgttccga actgtacaaa tacaaagtta 3300
tcgaactgaa accgctgggt atcgctccga ccggtgctaa gcttaccggt caggctcgtc 3360
agctgctgtc ctccatcggt cagcagcagt ccaacctgct gcgtgctatc gaagctcagc 3420
agcacatgct gcagctgacc gtttggggta tcaaacagct gcagaccggt gttctggcta 3480
tcgaacgtta cctgaaagac cagcagctcg agatctggaa caacatgacc tggatggaat 3540
gggaccgtga aatctccaac tacaccgaca ccatctaccg tctgctggaa gactcccaga 3600
cccagcagga aaaaaacgaa aaagacctgc tggctctgga ctcttgaaa aacctgtggt 3660
cctgggtcga catctccaac tggctgtggt acatcaaaat cttcatcatg atcgttggtg 3720
gtctgatcgg tctgcgtatc atcttcgctg ttctgtccat ctgaggatcc agatctgctg 3780
tgcttcttag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgacctgg 3840
aaggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg cattgtctga 3900
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 3960
aagacaatag caggcatgct ggggatgcgg tgggctctat gggtagccag gtgctgaaga 4020
attgaccggg ttctcctggg gccagaaaga agcaggcaca tccccttctc tgtgacacac 4080
cctgtccacg cccctgggtc ttagttccag cccactcat aggacactca tagctcagga 4140
gggctccgcc ttcaatccca cccgctaaag tacttggagc ggtctctccc tccctcatca 4200
gccacacaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt 4260
aagtgcagag ggagagaaaa tgctccaac atgtgaggaa gtaatgagag aaatcataga 4320
attttaaggc catgatttaa ggccatcatg gccttaatct tccgcttctc cgctcactga 4380
ctcgctgcgc tcggtcggtc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4440
acggttatcc acagaatcag gggataacgc aggaagaaac atgtgagcaa aaggccagca 4500
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccataggc tccgcccccc 4560
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccgga caggactata 4620
aagataccag gcgtttcccc ctggaagctc cctcgctgcgc tctcctgttc cgacctgcc 4680
gcttaccgga tacctgtccg cttttctccc ttccgggaagc gtggcgcttt ctcatagctc 4740
acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 4800
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaacc 4860
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4920
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 4980
aacagtatbt ggtatctgcg ctctgctgaa gccagtacc ttccgaaaaa gagtggtag 5040
ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt ttttttgttt gcaagcagca 5100
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5160
cgctcagtggt aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 5220
cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 5280
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5340
tctatttctg tcatccatag ttgctgact cggggggggg gggcgctgag gtctgcctcg 5400
tgaagaaggt gttgctgact cataccaggc ctgaatcgcc ccatcatcca gccagaaagt 5460
gaggagacca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5520
ttgctttgcc acggaacggt ctgcgttgct ggggaagatgc gtgatctgat ccttcaactc 5580
agcaaaagt cgatttatcc aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc 5640
cagtgttaca accaatlaac caattctgat tagaaaaact catcgagcat caaatgaaac 5700
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5760
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggctctgcg 5820
attccgactc gtccaacatc aatacaacct attaatctcc cctcgctcaa aataaggtta 5880
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5940
atcttcttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 6000
tcaaccaaac cgttatctat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg 6060
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6120
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccc 6180
gggatcgag tggtagtaaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6240
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6300
gcaacgtac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatataat 6360
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6420
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccc ttgaatatgg 6480

```

```

ctcataaacac cccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6540
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttccccc 6600
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6660
tgtattttaga aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6720
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6780
ccctttcgtc                                     6790

```

&lt;210&gt; 143

&lt;211&gt; 6658

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 143

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcctg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggatct ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160
cgcaggaaat cgttctggaa aacggtaccg aaaacttcaa catgtggaaa aacgacatgg 2220

```



ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagcatcaac	tgcaacgctt	2340
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2400
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggtccgt	2460
gcaacaacgt	ttccaccggt	cagtgcaccc	atgggtatcaa	accggttggt	tccacccagc	2520
tgctgctgaa	cgggtccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2580
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgaccccgtc	2640
cgaacaacaa	cacccgtaaa	tccatgcgta	tcgggtccggg	tcagaccttc	tacgtaccg	2700
gtgacatcat	cgggtgacatc	cgtcaggctt	actgcaacat	ctccggttcc	aaatggaacg	2760
aaaccctgaa	acgtgtttaa	gaaaaactgc	aggaaaacta	caacaacaac	aaaaccatca	2820
aattcgctcc	gtcctccggt	ggtgacctgg	aaatcaccac	ccactccttc	aactgccgtg	2880
gtgaattctt	ctactgcaac	accacccgtc	tgttcaacaa	caacgctacc	gaagacgaaa	2940
ccatcacctt	gccgtgccgt	atcaaacaga	tcatacaacat	gtggcagggg	gttgggtcgtg	3000
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	3060
tgctgggtcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3120
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3180
cgtggttat	cgtccgacc	ggtgctaagc	ttaccgttca	ggctcgtcag	ctgctgtcct	3240
ccatcgttca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3300
agctgaccgt	ttgggggtatc	aaacagctgc	agaccctgtg	tctggctatc	gaacgttacc	3360
tgaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3420
tctccaacta	caccgacacc	atctaccgtc	tgctggaaga	ctcccagacc	cagcaggaaa	3480
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtcc	tggttcgaca	3540
tctccaactg	gctgtggtac	atcaaaatct	tcatacatgat	cgttgggtgg	ctgatcggtc	3600
tgcgtatcat	cttcgctggt	ctgtccatct	gaggatccag	atctgctgtg	ccttctagtt	3660
gccagccatc	tgttggtttgc	ccctcccccg	tgcttccctt	gacctggaa	ggtgccactc	3720
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3780
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3840
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccaggt	gctgaagaat	tgaccgggtt	3900
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	3960
cctgggtctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	4020
caatcccacc	cgttaaagta	cttggagcgg	tctctccctc	cctcatcagc	ccaccaaacc	4080
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	4140
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	4200
tgattttaagg	ccatcatggc	cttaatcttc	cgttccctcg	ctcactgact	cgctgcgctc	4260
ggtcggtcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtataac	ggttatccac	4320
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4380
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgcacctctg	acgagcatca	4440
caaaaatcga	cgtcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	4500
gtttccccct	ggaagctccc	tctgtcgctc	tctgttcccg	accctgcgcg	ttaccggata	4560
cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	4620
tctcagttcg	gtgtagggtcg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	4680
gcccagccgc	tgcgccttat	ccggtaaacta	tcgtcttgag	tccaacccgg	taagacacga	4740
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgagggt	atgtaggcgg	4800
tgtacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	4860
tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	4920
caaacaaacc	accgctggta	gcgggtgggtt	ttttgtttgc	aagcagcaga	ttacgcgcag	4980
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	5040
cgaaaactca	cgttaaggga	ttttggtcat	gagattatca	aaaaggatct	tcacctagat	5100
ccttttaaat	taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	5160
tgacagttac	caatgcttaa	tcagtgagge	acctatctca	gcgatctgtc	tatttcgttc	5220
atccatagtt	gcctgactcg	gggggggggg	gcgctgaggt	ctgcctcgtg	aagaagggtg	5280
tgctgactca	taccaggcct	gaatcgcccc	atcatccagc	cagaaagtga	gggagccacg	5340
gttgatgaga	gctttgttgt	aggtggacca	gttgggtgatt	ttgaactttt	gctttgccac	5400
ggaacgggtct	gcgttgctcg	gaagatgcgt	gatctgatcc	ttcaactcag	caaaagtctg	5460
atttattcaa	caaagccgcc	gtcccgtaaa	gtcagcgtaa	tgctctgcca	gtgttacaac	5520

```

caattaacca attctgatta gaaaaactca tctgagcatca aatgaaactg caattttatc 5580
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5640
tcaccgaggc agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5700
ccaacatcaa tacaacctat taattttccc tctgcaaaaa taagggtatc aagtgagaaa 5760
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5820
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac cactcgcac 5880
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgcgtgtt aaaaggacaa 5940
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac cactcgcac 6000
tcacctgaat caggatattc ttctaatacc tggaaatgctg ttttcccggg gatcgcagtg 6060
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6120
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6180
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6240
gcacctgatt gcccgcacatt atcgcgagcc cattttatacc catataaatc agcatccatg 6300
ttggaattta atcgcggcct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6360
cttgatttac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6420
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc cccccattat 6480
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6540
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6600
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6658

```

&lt;210&gt; 144

&lt;211&gt; 6577

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 144

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgc tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttgggtcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgccttggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatgggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440

```



catgggctct tctccggtag cggcggagct tccacatccg agccctggte ccatgcctcc 1500  
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgacag gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcgggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctcttcc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920  
ccatgcgtgt tcgtgggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980  
gtttctggat gatcatcctc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040  
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100  
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160  
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220  
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280  
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340  
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2400  
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggg aaaggctccg 2460  
gcaacaacgt ttccaccggt cagtgcaccc atgggtatcaa accggttgtt tccaccagc 2520  
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580  
acaacgttaa aaccatcctc gttcacctga acaaatccgt tgaaatcggt tgcaccctg 2640  
ctagcgctta ctgcaacatc tccggttcca aatggaacga aacctgaaa cgtgttaaag 2700  
aaaaactgca ggaaaactac aacaacaaca aaaccatcaa attcgctccg tcctccgggt 2760  
gtgacctgga aatcaccacc cactccttca actgcctggg tgaattcttc tactgcaaca 2820  
ccaccgtct gttcaacaac aacgctaccg aagacgaaac catcacctg ccgtgccgta 2880  
tcaaacagat catcaacatg tggcagggtg ttggctcgtc tatgtacgt ccgccgatcg 2940  
ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctgggtcgt gacggtgggtg 3000  
aagacaacaa aaccgaagaa atcttccgtc cgggtgggtg taacatgaaa gacaactggc 3060  
gttccgaact gtacaaatac aaagtattcg aactgaaacc gctgggtatc gctccgaccg 3120  
gtgctaagct taccgttcag gctcgtcagc tgctgtcctc catcgttcag cagcagcca 3180  
acctgctgcg tgctatcgaa gctcagcagc acatgctgca gctgaccgtt tggggatatca 3240  
aacagctgca gaccgtgtt ctggctatcg aacgttacct gaaagaccag cagctcgaga 3300  
tctggaacaa catgacctgg atggaatggg accgtgaaat ctccaactac accgacacca 3360  
tctaccgtct gctggaagac tcccagaccc agcaggaaaa aaacgaaaaa gacctgctgg 3420  
ctctggactc ctggaaaaac ctgtgggtcct ggttcgacat ctccaactgg ctgtggtaca 3480  
tcaaaatctt catcatgac gttggtgggtc tgatcggtct gcgtatcctc ttcgctgttc 3540  
tgtccatctg aggatccaga tctgctgtgc cttctagttg ccagccatct gttgtttgcc 3600  
cctccccgt gccttccctg accctggaag gtgccactcc cactgtcctt tcctaataaa 3660  
atgaggaaat tgcctcgcct tgtctgagta ggtgtcatc tattctgggg ggtgggggtg 3720  
ggcaggacag caagggggag gattgggaag acaatagcag gcctgctggg gatgcgggtg 3780  
gctctatggg taccaggtg ctgaagaatt gaccgggttc ctctggggcc agaaagaagc 3840  
aggcacatcc ccttctctgt gacacaccct gtccacgccc ctgggttctta gttccagccc 3900  
cactcatagg aactcatag ctccaggagg ctccgccttc aatcccaccc gctaaagtac 3960  
ttggagcggg ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg 4020  
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4080  
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4140  
ttaatcttcc gcttccctgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4200  
ggatcatgct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4260  
aaagaacatg tgagcaaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct 4320  
ggcgttttcc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4380  
gaggtggcga aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4440  
cgtgcgctct cctgttccga ccctgccgct taccggatc ctgtccgct tcttcccttc 4500  
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctgagttcgg tbtaggctct 4560  
tcgctccaag ctgggctgtg tgcacgaacc cccggttcag cccgaccgct gcgccttate 4620  
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4680  
cactggtaac aggattagca gagcgaggta tbtaggcggg gctacagagt tcttgaagtg 4740

```

gtggcctaac tacggctaca ctagaagaac agtattttggt atctgcgctc tgctgaagcc 4800
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4860
cgggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 4920
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 4980
tttggtcacg agattatcaa aaaggatctt cacctagatc cttttaaaatt aaaaatgaag 5040
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5100
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5160
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5220
aatcgcccca tcatccagcc agaaagtga ggagccacgg ttgatgagag ctttgttgta 5280
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacgggtctg cgttgctcgg 5340
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 5400
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaacca ttctgattag 5460
aaaaactcat cgagcatcaa atgaaactgc aattttattca tatcaggatt atcaatacca 5520
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5580
atggcaagat cctggatcgc gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5640
aatttccctc cgtcaaaaat aagggttatca agtgagaaat caccatgagt gacgactgaa 5700
tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5760
ttacgctcgt catcaaaatc actcgcacac accaaaccgt tattcattcg tgattgcgcc 5820
tgagcgagac gaaatacgcg atcgtctgta aaaggacaat tacaacagg aatcgaatgc 5880
aaccggcgca ggaacactgc cagcgcacac acaatathtt cacctgaatc aggatattct 5940
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcacatca 6000
ggagtacgga taaaatgctt gatggctcga agaggcataa attccgtcag ccagtttagt 6060
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6120
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6180
tcgcgagccc atttataccc atataaatca gcatccatgt tgggaatttaa tcgcggcctc 6240
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6300
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6360
ttttgagaca caacgtggct tcccccccc cccattattt gaagcattta tcagggttat 6420
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6480
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6540
acctataaaa ataggcgatb caccaggccc tttcgtc 6577

```

&lt;210&gt; 145

&lt;211&gt; 6538

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 145

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcatatgac ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720

```

catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgct	agatcgccct	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgaag	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggtc	ccattttatta	tttaciaaatt	cacatataca	acaacgccgt	cccccggtgcc	1380
cgcagttttt	attaaacata	gcgtggggtc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggtgag	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtggtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgttgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcttttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatacgcg	ccgctctaga	gatatacgcca	1920
ccatgcgtgt	tctgtgtatc	ccgcgttaact	ggccgcagtg	gtggatgtgg	ggatatcctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtt	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagcatcaac	tgcaacgctt	2340
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2400
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggctccgt	2460
gcaacaacgt	ttccaccggt	cagtgcaccc	atggtatcaa	accggttgtt	tccaccagc	2520
tgtgtctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2580
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaatcgtt	tgcacccgtg	2640
ctagcgctta	ctgcaacatc	tccggttcca	aatggaacga	aacctgaaa	cgtgttaaag	2700
aaaaactgca	ggaaaactac	aacaacaaca	aaaccatcaa	attcgctccg	tcctccggtg	2760
gtgacctgga	aatcaccacc	cactccttca	actgcctggg	tgaattcttc	tactgcaaca	2820
ccgctagcct	gccgtgccgt	atcaaacaga	tcatacaatc	gtggcagggg	gttggtcgtg	2880
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	2940
tgtgtgttcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3000
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3060
cgtgtgggtat	cgtcccgacc	gggtgctaagc	ttaccgttca	ggctcgtcag	ctgctgtcct	3120
ccatcgttca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3180
agctgaccgt	ttgggttatc	aaacagctgc	agaccctgtg	tctggctatc	gaacgttacc	3240
tgaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3300
tctccaacta	caccgacacc	atctaccgtc	tgttgaaga	ctcccagacc	cagcaggaaa	3360
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtcc	tgggtcgcaca	3420
tctccaactg	gctgtggtac	atcaaaaatc	tcatacatgat	cgttgggtgg	ctgatcgggtc	3480
tgcgtatcat	cttcgctgtt	ctgtccatct	gaggatccag	atctgctgtg	ccttctagtt	3540
gccagccatc	tgttgtttgc	ccctcccccg	tgccttctct	gacctggaa	ggtgccactc	3600
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3660
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3720
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtacccaggt	gctgaagaat	tgacccgggt	3780
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	3840
cctggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	3900
caatcccacc	cgttaaagta	cttgagcggt	tctctccctc	cctcatcagc	ccaccaaacc	3960
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctabtaa	gtgcagaggg	4020

agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat ttttaaggcca 4080  
tgattttaagg ccatcatggc cttaatcttc cgcttcctcg ctactgact cgctgcgctc 4140  
ggtcgttcgg ctgcggcgag cggatcagc tcaactcaaag gcggtaatac ggttatccac 4200  
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4260  
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 4320  
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 4380  
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 4440  
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4500  
tctcagttcg gtgtaggtcg ttgcgtccaa gctgggctgt gtgcacgaac ccccggttca 4560  
gcccgaaccg tgcgccttat ccggtacta tcgtcttgag tccaacccgg taagacacga 4620  
cttatcgcca ctggcagcag ccaactggtaa caggattagc agagcgaggt atgtaggcgg 4680  
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4740  
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4800  
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 4860  
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg ctcagtggaa 4920  
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 4980  
cctttttaat taaaaatgaa gtttttaaatc aatctaaagt atatatgagt aaacttggtc 5040  
tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc tatttcgttc 5100  
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5160  
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5220  
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5280  
ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 5340  
atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5400  
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caattttattc 5460  
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5520  
tcaccgaggc agttccatag gatggcaaga tcctgggtatc ggtctgcgat tccgactcgt 5580  
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5640  
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5700  
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaaccg 5760  
ttattcatte gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5820  
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5880  
tcacctgaat caggatatte ttctaatacc tggaatgctg ttttcccggt gatcgagtg 5940  
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggctcg aagaggcata 6000  
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6060  
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6120  
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6180  
ttggaattta atcgcgccct cgagcaagac gtttcccgtt gaatatggct cataacaccc 6240  
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6300  
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccatatt 6360  
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6420  
aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6480  
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6538

&lt;210&gt; 146

&lt;211&gt; 6718

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 146

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60



cagcttgtct	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcggtg	tcggggctgg	cttaactatg	cggcatacaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaattg	420
cccgccctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tgcccgccct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgcg	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgga	780
cgtcaatggg	agtttggttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgaggagggt	atataagcag	900
agctcgttta	gtgaaccgct	agatcgccctg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	accccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tatttttaca	1320
ggatgggggtc	ccatttatta	tttaciaaatt	cacatataca	acaacgccgt	ccccgtgcc	1380
cgcagttttt	attaaacata	gcgtgggata	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgcctcg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcgggtgc	tgttaacggg	ggagggcag	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tcgtgggtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggatcctgg	1980
gtttctggat	gatcatcatc	tgcctgtgtg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gctccttcgc	atgcatcaac	tgcaacgctt	2400
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2460
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggctccg	2520
gcaacaacgt	ttccaccggt	cagtgcaccc	atgggtatcaa	accggttggt	tccaccacgc	2580
tgctgctgaa	cggttccctg	gctgaaaaag	aaatcatcat	ccgttccgaa	aacctgaccg	2640
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgcacccgtc	2700
cgaacaacaa	caccgcgttaa	tccatgcgta	tccgtccggg	tcagaccttc	tacgctaccg	2760
gtgacatcat	cggtgacatc	cgtcaggctt	actgcaacat	ctccgggttc	aaatggaacg	2820
aaaccctgaa	acgtgttaaa	gaaaaactgc	aggaaaacta	caacaacaac	aaaaccatca	2880
aattcgctcc	gtcctccggt	ggtgacctgg	aaatcaccac	ccactccttc	aactgccgtg	2940
gtgaattctt	ctactgcaac	accacccgtc	tgttcaacaa	caacgctacc	gaagacgaaa	3000
ccatcacctt	gccgtgccgt	atcaaacaga	tcacatcaac	gtggcagggt	gttggctcgtg	3060
ctatgtacgc	tccgccgatc	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	3120
tgctgggttcg	tgacgggtgg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3180
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3240
cgtcgggtat	cgtccgacc	ggtgctaagc	ttaccgttca	ggctcgtcag	ctgctgtcct	3300
ccatcggtca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3360



```

agctgaccgt ttgggggtatc aaacagctgc agacccgtgt tctggctatc gaacgttacc 3420
tgaaagacca gcagctcgag atctggaaca acatgacctg gatggaatgg gaccgtgaaa 3480
tctccaacta caccgacacc atctaccgtc tgctggaaga ctcccagacc cagcaggaaa 3540
aaaacgaaaa agacctgctg gctctggact cctggaaaaa cctgtggtcc tggttcgaca 3600
tctccaactg gctgtggtac atcaaaatct tcatcatgat cgttggtggt ctgatcggtc 3660
tgcgatatcat ctctcgctgtt ctgtccatct gaggatccag atctgctgtg ccttctagtt 3720
gccagccatc tgttgtttgc cctcccccg tgccttcctt gacctggaa ggtgccactc 3780
ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 3840
ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 3900
ggcatgctgg ggatgcggtg ggctctatgg gtaccaggt gctgaagaat tgaccgggtt 3960
cctcctgggc cagaaagaag caggcacatc cccttctctg tgacacaccc tgtccacgcc 4020
cctgggttctt agttccagcc ccactcatag gacactcata gctcaggagg gctccgcctt 4080
caatcccacc cgctaaagta cttggagcgg tctctccctc cctcatcagc ccaccaaacc 4140
aaacctagcc tccaagagt ggaagaaatt aaagcaagat aggctattaa gtgcagaggg 4200
agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat ttttaaggcca 4260
tgattttaagg ccatcatggc cttaatcttc cgcttccctc ctcactgact cgctgcgctc 4320
ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaa gcggttaatac ggttatccac 4380
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4440
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 4500
caaaaatcga cgctcaagtc agagggtggc aaaccgcaca ggactataaa gataccaggc 4560
gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 4620
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 4680
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac ccccgttca 4740
gcccgaccgc tgcgccttat ccggtacta tctgtttgag tccaaccgg taagacacga 4800
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4860
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 4920
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 4980
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 5040
aaaaaaagga tctcaagaag atcctttgat cttttctacg ggtctgacg ctcagtggaa 5100
cgaaaactca cgttaaggga ttttggtcac gagattatca aaaaggatct tcacctagat 5160
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 5220
tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc tatttcgttc 5280
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5340
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5400
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5460
ggaacggtct gcgttgtcgg gaagatgcgt gatctgatcc ttcaactcag caaaagttcg 5520
atttattcaa caaagccgcc gtcccgtaa gtcagcgtaa tgctctgcca gtgttacaac 5580
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caatttatc 5640
atatcaggat tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac 5700
tcaccgaggc agttccatag gatggcaaga tcctggatc ggtctgcgat tccgactcgt 5760
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa 5820
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5880
acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac aaccaaaccg 5940
ttattcatte gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 6000
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatat 6060
tcacctgaat caggatatte ttctaatacc tggaatgctg ttttcccggg gatcgagtg 6120
gtgagtaacc atgcatcacc aggagtacgg ataaaatgct tgatggtcgg aagaggcata 6180
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6240
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6300
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6360
ttggaattta atcgcggcct cgagcaagac gtttcccggt gaatatggct cataacaccc 6420
cttgtattac tgtttatgta agcagacagt tttattgttc atgatgatat atttttatct 6480
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccatat 6540
tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 6600
aataaaciaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6660

```

accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6718

<210> 147

<211> 6637

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 147

tgcgcgcttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360  
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480  
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600  
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660  
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
catcaatggg cgtggatagc ggtttgactc acggggatbt ccaagtctcc accccattga 780  
cgtcaatggg agtttgtttt ggcacccaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900  
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960  
tagaagacac cgggaccgat ccagccctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgcct ccttatgcta 1140  
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380  
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500  
agcggctcat ggctcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gltgcgggtc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920  
ccatgcgtgt tcgtgggtat ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980  
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040  
gtgttccggg ttggaccgac gctaaaacca ccctgtcttg cgcttccgac accaaagcct 2100  
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gaccogaacc 2160  
cgcaggaaat cgttctggaa aacggttacg aaaacttcaa catgtggaaa aacgacatgg 2220  
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280  
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340  
ccaacgacat gaacaaagaa atccgtaact gctccttcgc atgcatcaac tgcaacgctt 2400  
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2460  
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggg aaagggtccg 2520

```

gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttggt tccaccacgc 2580
tgctgctgaa cgggtccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcggt tgcaccctg 2700
ctagecgctta ctgcaacatc tccggttcca aatggaacga aacctgaaa cgtgttaaag 2760
aaaaactgca ggaaaactac aacaacaaca aaaccatcaa attcgctccg tccctccggtg 2820
gtgacctgga aatcaccacc cactccttca actgccgtgg tgaattcttc tactgcaaca 2880
ccaccctgtt gttcaacaac aacgctaccg aagacgaaac catcaccctg ccgtgccgta 2940
tcaaacagat catcaacatg tggcaggggt ttggtcgtgc tatgtacgct ccgccgatcg 3000
ctggtaacat cacctgcaaa tccaacatca ccggtctgct gctgggtcgt gacgggtgggt 3060
aagacaacaa aaccgaagaa atcttccgtc cgggtgggtg taacatgaaa gacaactggc 3120
gttccgaact gtacaaatac aaagtatatc aactgaaacc gctgggtatc gctccgaccg 3180
gtgctaagct taccgttcag gctcgtcagc tgctgtcctc catcggttcag cagcagtcca 3240
acctgctgct tgctatcgaa gctcagcagc acatgctgca gctgaccgtt tgggggtatca 3300
aacagctgca gaccctgtgt ctggctatcg aacgttacct gaaagaccag cagctcgaga 3360
tctggaacaa catgacctgg atggaatggg accgtgaaat ctccaactac accgacacca 3420
tctaccgtct gctggaagac tccagacccc agcaggaaaa aaacgaaaaa gacctgctgg 3480
ctctggactc ctggaaaaac ctgtgggtcct gggtcgacat ctccaactgg ctgtgggtaca 3540
tcaaaatctt catcatgata gttgggtggc tgatcggtct gcgtatcatc ttcgctgttc 3600
tgtccatctg aggatccaga tctgctgtgc cttctagttg ccagccatct gttgtttgcc 3660
cctccccctg gccttccctg accctggaag gtgccactcc cactgtcctt tccataataa 3720
atgaggaaat tgcategcac tgtctgagta ggtgtcattc tattctgggg ggtgggggtg 3780
ggcaggacag caagggggag gattgggaag acaatagcag gcattgctgg gatgcgggtg 3840
gctctatggg taccaggtg ctgaagaatt gaccgggttc ctctggggcc agaaagaagc 3900
aggcacatcc ccttctctgt gacacacccc gtccacgccc ctggttctta gttccagccc 3960
cactcatagg aactcatag ctccaggagg ctccgccttc aatccccccc gctaaagtac 4020
ttggagcggt ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg 4080
gaagaaatta aagcaagata ggctattaag tgcagaggga gagaaaatgc ctccaacatg 4140
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4200
ttaatcttcc gcttccctgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4260
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4320
aaagaacatg tgagcaaaag gccagcaaaa ggccagggaac cgtaaaaagg ccgcgttgct 4380
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440
gaggtggcga aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4560
gggaagcgtg gcgtttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 4620
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc 4680
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc 4740
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 4800
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 4860
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920
cgggtggttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga 4980
tcctttgate ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040
tttgggtcat agattatcaa aaaggatctt cacctagatc cttttaaat taaaaatgaag 5100
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 5160
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagtgt cctgactcgg 5220
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5280
aatcgcccca tcatccagcc agaaagttag ggagccacgg ttgatgagag ctttgttgta 5340
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggctct cgttgtcggg 5400
aagatgcgtg atctgatect tcaactcagc aaaagttcga ttattcaac aaagccgccg 5460
tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5520
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640
atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5700
aatttcccct cgtcaaaaat aagggtatca agtgagaaat caccatgagt gacgactgaa 5760
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820

```



```

ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5880
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaaacagg aatcgaatgc 5940
aaccggcgca ggaacactgc cagcgcatca acaatatatt cacctgaatc aggatattct 6000
tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca tgcacatca 6060
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag ccagtttagt 6120
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6240
tcgcgagccc atttataccc atataaatca gcatccatgt tgggaatttaa tcgcggcctc 6300
gagcaagacg tttcccggtg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420
ttttgagaca caacgtggct tcccccccc cccattattt gaagcattta tcagggttat 6480
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600
acctataaaa ataggcgat cagaggccc tttcgtc 6637

```

&lt;210&gt; 148

&lt;211&gt; 6598

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 148

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aagaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tgcccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtcctc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggatgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcctg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttcgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgcctcg cagtcctctg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcctttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tcgtggtatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgcttccgac	accaaagcct	2100
acgaccgtga	agttcacaac	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	taccttcaaa	aacaacgtta	2340
ccaacgacat	gaacaaagaa	atccgtaact	gtcccttcgc	atgcatcaac	tgcaacgctt	2400
ccaccatcac	ccaggcttgc	ccgaaagtta	acttcgaccc	gatcccgatc	cactactgcg	2460
ctccggctgg	ttacgctatc	ctgaaatgca	acaacaaaac	cttctccggt	aaaggtccgt	2520
gcaacaacgt	ttccaccggt	cagtgcaccc	atggtatcaa	accggttggt	tccaccagc	2580
tgctgctgaa	cgggttccctg	gctgaaaaag	aatcatcat	ccgttccgaa	aacctgaccg	2640
acaacgttaa	aaccatcatc	gttcacctga	acaaatccgt	tgaaatcggt	tgaccccggt	2700
ctagcgctta	ctgcaacatc	tccggttcca	aatggaacga	aacctgaaa	cgtgttaaag	2760
aaaaactgca	ggaaaactac	aacaacaaca	aaaccatcaa	atcgctccg	tcctccggtg	2820
gtgacctgga	aatcaccacc	cactccttca	actgcgctgg	tgaattcttc	tactgcaaca	2880
ccgctagcct	gccgtgccgt	atcaaacaga	tcaccaacat	gtggcagggt	gttggtcgtg	2940
ctatgtacgc	tccgccgac	gctggtaaca	tcacctgcaa	atccaacatc	accggtctgc	3000
tgctgggttcg	tgacggtggg	gaagacaaca	aaaccgaaga	aatcttccgt	ccgggtgggtg	3060
gtaacatgaa	agacaactgg	cgttccgaac	tgtacaaata	caaagttatc	gaactgaaac	3120
cgctgggtat	cgctccgacc	gggtgctaagc	ttaccgttca	ggctcgtcag	ctgctgtcct	3180
ccatcggtca	gcagcagtc	aacctgctgc	gtgctatcga	agctcagcag	cacatgctgc	3240
agctgaccgt	ttggggatc	aaacagctgc	agaccctgtg	tctggctatc	gaacgttacc	3300
tgaaagacca	gcagctcgag	atctggaaca	acatgacctg	gatggaatgg	gaccgtgaaa	3360
tctccaacta	caccgacacc	atctaccgtc	tgctggaaga	ctcccagacc	cagcaggaaa	3420
aaaacgaaaa	agacctgctg	gctctggact	cctggaaaaa	cctgtgggtcc	tggttcgaca	3480
tctccaactg	gctgtggtac	atcaaaatct	tcacatgat	cgttgggtgg	ctgatcggtc	3540
tgcgtatcat	cttcgctggt	ctgtccatct	gaggatccag	atctgctgtg	ccttctagtt	3600
gccagccatc	tgttgtttgc	ccctcccccg	tgccttccct	gacctggaa	gggtgccactc	3660
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	3720
ctattctggg	gggtggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	3780
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtaccagagt	gctgaagaat	tgaccgggtt	3840
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacaccc	tgtccacgcc	3900
cctgggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	3960
caatcccacc	cgctaaagta	cttgagcgcg	tcttccctc	cctcatcagc	ccaccaaacc	4020
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	4080
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	4140
tgattttaagg	ccatcatggc	cttaatcttc	cgttccctcg	ctcactgact	cgctgcgctc	4200
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaa	gcggtaatat	ggttatccac	4260
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	4320
ccgtaaaaag	gccgcgttgc	tggcggtttt	ccataggctc	cgccccctg	acgagcatca	4380
caaaaatcga	cgctcaagtc	agagggtggc	aaacccgaca	ggactataaa	gataccaggc	4440
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	4500
cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	4560
tctcagttcg	gtgtaggtcg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	4620
gcccagaccg	tgcgccttat	ccggtaacta	tcgtcttgag	tccaacccgg	taagacacga	4680
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	4740
tgctacagag	ttcttgaagt	gggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	4800
tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggttagct	cttgatccgg	4860
caaacaaacc	accgctggta	gcgggtgggtt	ttttgtttgc	aagcagcaga	ttacgcgcag	4920
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	4980
cgaaaactca	cgttaaggga	ttttgggtcat	gagattatca	aaaaggatct	tcacctagat	5040



```

cctttttaa at taaaaatgaa gtttttaa at aatctaaagt atatatgagt aaacttggtc 5100
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 5160
atccatagtt gcctgactcg gggggggggg gcgctgaggt ctgcctcgtg aagaagggtg 5220
tgctgactca taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg 5280
gttgatgaga gctttgttgt aggtggacca gttggtgatt ttgaactttt gctttgccac 5340
ggaacggtct gcgttgctcg gaagatgcgt gatctgatcc ttcaactcag caaaagtctg 5400
atltatltcaa caaagccgcc gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac 5460
caattaacca attctgatta gaaaaactca tcgagcatca aatgaaactg caattttattc 5520
atatcaggat tatcaatacc atatttttga aaaagccggt tctgtaatga aggagaaaac 5580
tcaccgaggc agttccatag gatggcaaga tcctgggtatc ggtctgcatc tccgactcgt 5640
ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa taagggtatc aagtgagaaa 5700
tcaccatgag tgacgactga atccggtgag aatggcaaaa gcttatgcat ttctttccag 5760
acttggtcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaacccg 5820
ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa 5880
ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcatc aacaatattt 5940
tcacctgaat caggatattc ttctaatacc tggaaatgctg ttttcccggt gatcgcatg 6000
gtgagtaacc atgcatcatc aggagtacgg ataaaatgct tgatggctcg aagaggcata 6060
aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct 6120
ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc 6180
gcacctgatt gcccgacatt atcgcgagcc catttatacc catataaatc agcatccatg 6240
ttggaattta atcgcgccct cgagcaagac gtttcccggt gaatatggct cataaacacc 6300
cttgatttac tgtttatgta agcagacagt tttattgttc atgatgat atttttatct 6360
tgtgcaatgt aacatcagag attttgagac acaacgtggc tttccccccc ccccatatt 6420
tgaagcattt atcagggtta ttgtctcatg agcggatata tttttgaatg tttttgaaa 6480
aataa caaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 6540
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 6598

```

&lt;210&gt; 149

&lt;211&gt; 6679

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 149

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgc tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatagcgt gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcattg gaacgcggat 1020

```

```

tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcc aactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttaaaatt cacatataca acaacgcggt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggtagc tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920
ccatgcgtgt tctgtgtatc ccgcgtact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtact gctccttcgc atgcatcaac tgcaacgctt 2400
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2460
ctccggctgg ttacgtatc ctgaaatgca acaacaaaac cttctccggt aaaggctcgt 2520
gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttgtt tccaccagc 2580
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2640
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcgtt tgcaccgctc 2700
cgaacaacaa caccgtaaa tccatgcgta tccgtccggg tcagaccttc tacgtaccg 2760
gtgacatcat cggtgacatc cgtcaggctt actgcaacat ctccggttcc aaatggaacg 2820
aaacctgaa acgtgttaaa gaaaaactgc aggaaaacta caacaacaac aaaacctatc 2880
aatctgctcc gtctccggtt ggtgacctgg aaatcaccac ccactccttc aactgccgtg 2940
gtgaattctt ctactgcaac accgctagcc tgccgtgccc tatcaaacag atcatcaaca 3000
tgtggcaggg tgttggtcgt gctatgtacg ctccgccgat cgctggtaac atcacctgca 3060
aatccaacat caccggtctg ctgctggttc gtgacggtgg tgaagacaac aaaaccgaag 3120
aaatcttccg tccgggtggt ggtaacatga aagacaactg gcgttccgaa ctgtacaaat 3180
acaaagttat cgaactgaaa ccgctgggta tccgtccgac cggtgctaag cttaccgttc 3240
aggctcgtca gctgctgtcc tccatcgctt agcagcagtc caacctgctg cgtgctatcg 3300
aagctcagca gcacatgctg cagctgaccg tttggggtat caaacagctg cagaccgctg 3360
ttctggctat cgaacgttac ctgaaagacc agcagctcga gatctggaac aacatgacct 3420
ggatggaatg ggaccgtgaa atctccaact acaccgacac catctaccgt ctgctggaag 3480
actcccagac ccagcaggaa aaaaacgaaa aagacctgct ggctctggac tcctggaaaa 3540
acctgtggtc ctggttcgac atctccaact ggctgtggta catcaaaatc ttcacatga 3600
tcgttggtgg tctgatcggg ctgcgtatca tcttcgctgt tctgtccatc tgaggatcca 3660
gatctgctgt gccttctagt tgccagccat ctggtgtttg cccctcccc gtgccttctt 3720
tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc 3780
attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac agcaaggggg 3840
aggattggga agacaatagc aggcagtctg gggatgcggt gggctctatg ggtaccagc 3900
tgctgaagaa ttgaccgggt tctcctggg ccagaaagaa gcaggcacat ccccttctct 3960
gtgacacacc ctgtccacgc ccctgggtct tagttccagc cccactcata ggacactcat 4020
agctcaggag ggctccgct tcaatcccac ccgctaaagt acttgagcgc gtctctccct 4080
ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaagaaat taaagcaaga 4140
taggctatta agtgagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 4200
aatcatagaa ttttaaggcc atgatttaag gccatcatgg ccttaatctt ccgcttctc 4260
gctcactgac tccgtgcgct cggctcgttcg gctgcggcga gcggtatcag ctactcaaa 4320

```

```

ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4380
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 4440
ccgccccct gacgagcatc acaaaaatcg acgtcaagt cagagggtggc gaaacccgac 4500
aggactataa agataccagg cgtttccccc tggaagctcc ctctgctgct ctctgttcc 4560
gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc 4620
tcatagctca cgctgtagg atctcagttc ggtgtaggtc gttcgctcca agctgggctg 4680
tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga 4740
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag 4800
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta 4860
cactagaaga acagtatttg gtatctgcgc tctgtgaag ccagttacct tcggaaaaag 4920
agttggtagc tcttgatccg gcaaacaaac caccgctggg agcgggtgggt tttttgtttg 4980
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac 5040
ggggtctgac gctcagtggg acgaaaactc acgttaaggg attttgggtca tgagattatc 5100
aaaaaggatc ttcacctaga tctttttaa ttaaaaatga agttttaaat caatctaaag 5160
tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc 5220
agcgatctgt ctatttcgtt catccatagt tgcctgactc gggggggggg ggcgctgagg 5280
tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag 5340
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5400
tttgaacttt tgctttgcca cggaacggtc tgcgttgctg ggaagatgcg tgatctgac 5460
cttcaactca gcaaaagttc gatttatcca acaaagccgc cgtcccgtca agtcagcgtg 5520
atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc atcgagcatc 5580
aatgaaact gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5640
ttctgtaatg aaggagaaaa ctaccgagg cagttccata ggatggcaag atcctgggat 5700
cggctctgca ttcgactcg tccaacatca atacaaccta ttaatttccc ctctcaaaa 5760
ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga gaatggcaaa 5820
agcttatgca tttctttcca gacttggttc acaggccagc cattacgctc gtcataaaa 5880
tactcgcac caaccaaac gttattcatt cgtgattgct cctgagcgag acgaaatacg 5940
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 6000
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6060
gttttcccg ggatcgaggt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6120
ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgacct ctcatctgta 6180
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6240
ccatacaatc gatagattgt cgcacctgat tgcccgacat tatcgcgagc ccatttatac 6300
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccg 6360
tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattgtt 6420
catgatgata tatttttata ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6480
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggtac 6540
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa 6600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6660
atcacgaggc cctttcgtc
6679

```

&lt;210&gt; 150

&lt;211&gt; 6748

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 150

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggggcgc tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

```



```

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtct ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatatata acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgectcc 1500
agcggctcat ggtcgtcctg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagctctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtggatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gaccogaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccaccacc gaaatccgtg 2400
acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460
accgtaacaa ctccaacaac tccgaatata tctgatcaa ctgcaacgct tccaccatca 2520
cccaggcttg cccgaaagtt aacttcgacc cgatcccgat ccactactgc gctccggctg 2580
gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaagggtccg tgcaacaacg 2640
tttccacogt tcagtgcacc catggtatca aacgggttgt tttccaccag ctgctgctga 2700
acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760
aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcacctgt gctagcgtt 2820
actgcaacat ctccgggttc aaatggaacg aaacctgaa acgtgttaaa gaaaaactgc 2880
aggaaaacta caacaacaac aaaaccatca aattcgctcc gtcctccggt ggtgacctgg 2940
aaatcaccac ccactccttc aactgccgtg gtgaattctt ctactgcaac accaccgtc 3000
tgttcaacaa caacgctacc gaagacgaaa ccataccct gccgtgccgt atcaaacaga 3060
tcatcaacat gtggcagggt gttggtcgtg ctatgtacgc tccgccgatc gctggtaaca 3120
tcacctgcaa atccaacatc accggtctgc tgctggttcg tgacgggtgg gaagacaaca 3180
aaaccgaaga aatcttccgt ccgggtgggt gtaacatgaa agacaactgg cgttccgaac 3240
tgtacaaata caaagttatc gaactgaaac cgctgggtat cgctccgacc ggtgctaagc 3300
ttaccgttca ggctcgtcag ctgctgtcct ccacgttca gcagcagtc aacctgctgc 3360
gtgctatcga agctcagcag cacatgctgc agctgaccgt ttggggtatc aaacagctgc 3420
agacctgtgt tctggctatc gaacgttacc tgaaagacca gcagctcgag atctggaaca 3480

```

acatgacctg	gatggaatgg	gaccgtgaaa	tctccaacta	caccgacacc	atctaccgtc	3540
tgctggaaga	ctcccagacc	cagcaggaaa	aaaacgaaaa	agacctgctg	gctctggact	3600
cctggaaaaa	cctgtggtcc	tggttcgaca	tctccaactg	gctgtggtac	atcaaaatct	3660
tcatcatgat	cglttggtgg	ctgatcggtc	tgcgatcat	cttcgctgtt	ctgtccatct	3720
gaggatccag	atctgctgtg	ccttctagtt	gccagccatc	tggtgtttgc	ccctcccccg	3780
tgecttctt	gacctggaa	ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	3840
ttgcatcgca	ttgtctgagt	aggtgtcatt	ctattctggg	gggtgggggtg	gggcaggaca	3900
gcaaggggga	ggattgggaa	gacaatagca	ggcatgctgg	ggatgcgggtg	ggctctatgg	3960
gtacccaggt	gctgaagaat	tgacccgggt	cctcctgggc	cagaaagaag	caggcacatc	4020
cccttctctg	tgacacaccc	tgccacgcgc	cctggttctt	agttccagcc	ccactcatag	4080
gacactcata	gctcaggagg	gctccgcctt	caatcccacc	cgctaaagta	cttggagcgg	4140
tcttctccct	cctcatcagc	ccaccaaac	aaacctagcc	tccaagagt	ggaagaaatt	4200
aaagcaagat	aggctattaa	gtgcagaggg	agagaaaatg	cctccaacat	gtgaggaagt	4260
aatgagagaa	atcatagaat	tttaaggcca	tgatttaagg	ccatcatggc	cttaatcttc	4320
cgcttctctg	ctcactgact	cgctgcgctc	ggctggttcg	ctgcggcgag	cggtatcagc	4380
tcactcaaag	gcggtataac	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	4440
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcggtttt	4500
ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agaggtggcg	4560
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	4620
tctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	4680
ggcgctttct	catagctcac	gctgtaggt	tctcagttcg	gtgtaggtcg	ttcgctccaa	4740
gctgggctgt	gtgcacgaac	cccccgttca	gcccgaaccg	tgcgcccttat	ccggtaacta	4800
tcgtcttgag	tccaaccg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	4860
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	4920
ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	4980
cggaaaaaga	glttggtagct	cttgatccgg	caaacaacc	accgctggta	gcggtggttt	5040
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	5100
cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgttaaggga	ttttggtcat	5160
gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	gttttaaatc	5220
aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	5280
acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcg	gggggggggg	5340
gcgctgaggt	ctgcctcgtg	aagaagggtg	tgctgactca	taccaggcct	gaatcgcccc	5400
atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	gctttgttgt	aggtggacca	5460
gttggtgatt	ttgaactttt	gctttgccac	ggaacggtct	gcgttgctcg	gaagatgcgt	5520
gatctgatcc	ttcaactcag	caaaagttcg	atttattcaa	caaagccgcc	gtcccgtcaa	5580
gtcagcgtaa	tgctctgcca	gtgttacaac	caattaacca	attctgatta	gaaaaactca	5640
tcgagcatca	aatgaaactg	caatttatct	atatcaggat	tatcaatacc	atatttttga	5700
aaaagccgtt	tctgtaatga	aggagaaaac	tcaccgaggc	agttccatag	gatggcaaga	5760
tcctggatc	ggtctgcgat	tccgactcgt	ccaacatcaa	tacaacctat	taatttcccc	5820
tcgtcaaaaa	taaggttatc	aagtgagaaa	tcaccatgag	tgacgactga	atccggtgag	5880
aatggcaaaa	gcttatgcat	ttctttccag	acttgttcaa	caggccagcc	attacgctcg	5940
tcatacaaat	cactcgcac	aaccaaaccg	ttattcatct	gtgattgcgc	ctgagcgaga	6000
cgaaatacgc	gatcgctggt	aaaaggacaa	ttacaaacag	gaatcgaatg	caaccggcgc	6060
aggaacactg	ccagcgcac	aacaatatct	tcacctgaat	caggatatct	ttctaatacc	6120
tggaatgctg	ttttcccg	gatcgcgatg	gtgagtaacc	atgcatcatc	aggagtacgg	6180
ataaaatgct	tgatgggtcg	aagaggcata	aattccgtca	gccagtttag	tctgaccatc	6240
tcattctgtaa	catcattggc	aacgctacct	ttgccatggt	tcagaaacaa	ctctggcgca	6300
tcgggcttcc	catacaatcg	atagattgtc	gcacctgatt	gcccagacatt	atcgcgagcc	6360
catttatacc	catataaatc	agcatccatg	ttggaattta	atcgcgccct	cgagcaagac	6420
gtttcccggt	gaatatggct	cataacaccc	cttgtattac	tgtttatgta	agcagacagt	6480
tttattgttc	atgatgatat	atttttatct	tgtgcaatgt	aacatcagag	attttgagac	6540
acaacgtggc	tttccccccc	ccccatttat	tgaagcattt	atcaggggtta	ttgtctcatg	6600
agcggataca	tatttgaatg	tatttagaaa	aataaaca	taggggttcc	gcgcacattt	6660
ccccgaaaag	tgccacctga	cgtctaagaa	accattatta	tcatgacatt	aacctataaa	6720
aataggcgta	tcacgaggcc	ctttcgtc				6748



&lt;210&gt; 151

&lt;211&gt; 6709

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 151

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg ccatttgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaggt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatct ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac ttcccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa bgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggtctg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatthatta tttaaaaatt cacatataca acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgccg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtat ccgcgttaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttccggg ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gaccgaacc 2160
cgcaggaaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340
ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccaccacc gaaatccgtg 2400
acaaaaaaca gcagggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460
accgtaacaa ctccaacaac tccgaatata tctgatcaa ctgcaacgct tccaccatca 2520
cccaggcttg cccgaaagtt aacttcgacc cgatcccgat ccactactgc gctccggctg 2580

```

gttacgctat	cctgaaatgc	aacaacaaaa	ccttctccgg	taaagggtccg	tgcaacaacg	2640
tttccaccgt	tcagtgcacc	catgggtatca	aaccgggttg	ttccacccag	ctgctgctga	2700
acgggtccct	ggctgaaaaa	gaaatcatca	tccgttccga	aaacctgacc	gacaacgtta	2760
aaaccatcat	cggtcacctg	aacaaatccg	ttgaaatcgt	ttgcacccgt	gctagcgctt	2820
actgcaacat	ctccgggttc	aaatggaacg	aaaccctgaa	acgtgtttaa	gaaaaactgc	2880
aggaaaacta	caacaacaac	aaaaccatca	aattcgctcc	gtcctccggg	ggtgacctgg	2940
aaatcaccac	ccactccttc	aactgccgtg	gtgaattctt	ctactgcaac	accgctagcc	3000
tgccgtgccg	tatcaaacag	atcatcaaca	tgtggcaggg	tggttggtcgt	gctatgtacg	3060
ctccgccgat	cgctggtaac	atcacctgca	aatccaacat	caccgggtctg	ctgctgggttc	3120
gtgacgggtg	tgaagacaac	aaaaccgaag	aaatcttccg	tccgggtggg	ggtaacatga	3180
aagacaactg	gcgttccgaa	ctgtacaaat	acaaagttat	cgaactgaaa	ccgctgggta	3240
tcgctccgac	cgggtgctaag	cttaccgttc	aggctcgtca	gctgctgtcc	tccatcggtc	3300
agcagcagtc	caacctgctg	cgtgctatcg	aagctcagca	gcacatgctg	cagctgaccg	3360
tttgggggat	caaacagctg	cagacccgtg	ttctggctat	cgaacgttac	ctgaaagacc	3420
agcagctcga	gatctggaac	aacatgacct	ggatggaatg	ggaccgtgaa	atctccaact	3480
acaccgacac	catctaccgt	ctgctggaag	actcccagac	ccagcaggaa	aaaaacgaaa	3540
aagacctgct	ggctctggac	tcctggaaaa	acctgtgggtc	ctgggttcgac	atctccaact	3600
ggctgtggta	catcaaaatc	ttcatcatga	tcgttggtgg	tctgatcggt	ctgcgtatca	3660
tcttcgctgt	tctgtccatc	tgaggatcca	gatctgctgt	gccttctagt	tgccagccat	3720
ctgttggttg	cccctcccc	gtgccttcct	tgacctgga	agggtgccact	cccactgtcc	3780
tttcctaata	aaatgaggaa	attgcatcgc	attgtctgag	taggtgtcat	tctattcttg	3840
ggggtggggg	ggggcaggac	agcaaggggg	aggattggga	agacaatagc	aggcatgctg	3900
gggatgcggt	gggctctatg	ggtacccagg	tgctgaagaa	ttgaccgggt	tcctcctggg	3960
ccagaaagaa	gcaggcacat	ccccttctct	gtgacacacc	ctgtccacgc	ccctgggtct	4020
tagttccagc	cccactcata	ggacactcat	agctcaggag	ggctccgcct	tcaatcccac	4080
ccgctaaagt	acttggagcg	gtctctccct	ccctcatcag	cccaccaaac	caaacctagc	4140
ctccaagagt	gggaagaaat	taaagcaaga	taggtatta	agtgcagagg	gagagaaaat	4200
gcctccaaca	tgtgaggaag	taatgagaga	aatcatagaa	ttttaaggcc	atgatttaag	4260
gccatcatgg	ccttaatctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggtcgttcg	4320
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	4380
ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	4440
ggccgcgttg	ctggcggttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	4500
acgctcaagt	cagagggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	4560
tggaaagctc	ctcgtgcgct	ctcctgttcc	gacctgcgc	cttaccggat	acctgtccgc	4620
ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	cgtgttaggt	atctcagttc	4680
ggtgtaggte	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggttc	agcccgaccg	4740
ctgcgcctta	tccggtaact	atcgtcttga	gtccaaccgc	gtaagacacg	acttatcgcc	4800
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	4860
gttcttgaag	tgggtggccta	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	4920
tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	4980
caccgctggg	agcgggtggt	tttttggttg	caagcagcag	attacgcgca	gaaaaaaaagg	5040
atctcaagaa	gatcctttga	tcttttctac	ggggtctgac	gctcagtgga	acgaaaactc	5100
acgttaaggg	attttggtca	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	5160
ttaaaaatga	agtttttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	5220
ccaatgctta	atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	5280
tgcttgactc	gggggggggg	ggcgctgagg	tctgcctcgt	gaagaagggt	ttgctgactc	5340
ataccaggcc	tgaatcgccc	catcatccag	ccagaaagtg	agggagccac	ggttgatgag	5400
agctttggtg	taggtggacc	agttgggtgat	tttgaacttt	tgctttgcca	cggaacgggtc	5460
tgcggtgtcg	ggaagatgcg	tgatctgate	cttcaactca	gcaaaaagttc	gatttattca	5520
acaaagccgc	cgtcccgtca	agtcagcgta	atgctctgcc	agtgttataa	ccaattaacc	5580
aattctgatt	agaaaaactc	atcgagcatc	aatgaaact	gcaatttatt	catatcagga	5640
ttatcaatac	catatttttg	aaaaagccgt	ttctgtaatg	aaggagaaaa	ctcaccgagg	5700
cagttccata	ggatggcaag	atcctgggtat	cggctctgcga	ttccgactcg	tccaacatca	5760
atacaacctc	ttaatttccc	ctcgtcaaaa	ataaggttat	caagtgagaa	atcaccatga	5820
gtgacgactg	aatccgggtga	gaatggcaaa	agcttatgca	tttctttcca	gacttggtta	5880

```

acaggccagc cattacgctc gtcatacaaaa tcaactcgcat caaccaaacc gttatttcatt 5940
cgtgattgcg cctgagcgag acgaaatagc cgatcgctgt taaaaggaca attacaaaca 6000
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 6060
tcaggatatt cttctaatac ctggaatgct gttttcccgg ggatcgcagt ggtgagtaac 6120
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgctc 6180
agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 6240
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 6300
tgcccgacat tatcgcgagc ccattttatac ccatataaat cagcatccat gttggaattt 6360
aatcgcgggc tcgagcaaga cgtttcccgt tgaatatggc tcataacacc ccttgattta 6420
ctgtttatgt aagcagacag ttttattggt catgatgata tatttttatac ttgtgcaatg 6480
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 6540
tatcagggtt attgtctcat gagcggatac atattttgaat gtatttagaa aaataaacia 6600
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 6660
atcatgacat taacctataa aaataggcgt atcacgaggc cttttcgtc 6709

```

&lt;210&gt; 152

&lt;211&gt; 6790

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 152

```

tcgcgctt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt battaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacia 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgct agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```



gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga gatatcgcca 1920  
 ccatgcgtgt tcgtgggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980  
 gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040  
 gtgttccggt ttggaccgac gctaaaacca cctgttctg cgcttccgac accaaagcct 2100  
 acgaccgtga agttcacaac gtttgggcta cccacgcttg cgttccgacc gacccgaacc 2160  
 cgcaggaat cgttctggaa aacgttaccg aaaacttcaa catgtggaaa aacgacatgg 2220  
 ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280  
 aactgacccc gctgtgcgtt accctgcact gcaccaacgc taccttcaaa aacaacgtta 2340  
 ccaacgacat gaacaaagaa atccgtaact gctccttcaa caccaccacc gaaatccgtg 2400  
 acaaaaaaca gcaggggttac gctctgttct accgtccgga catcgttctg ctgaaagaaa 2460  
 accgtaacaa ctccaacaac tccgaataca tcctgatcaa ctgcaacgct tccaccatca 2520  
 cccaggttg ccgaaagtt aacttcgacc cgatcccgat ccactactgc gctccggctg 2580  
 gttacgctat cctgaaatgc aacaacaaaa ccttctccgg taaaggctcg tgcaacaacg 2640  
 tttccaccgt tcagtgcacc catggtatca aaccggtgtg ttccaccag ctgctgctga 2700  
 acggttccct ggctgaaaaa gaaatcatca tccgttccga aaacctgacc gacaacgtta 2760  
 aaaccatcat cgttcacctg aacaaatccg ttgaaatcgt ttgcaccctg ccgaacaaca 2820  
 acaccgtaa atccatgcgt atcgggtccg gtcagacctt ctacgctacc ggtgacatca 2880  
 tcggtgacat ccgtcaggct tactgcaaca tctccggttc caaatggaac gaaaccctga 2940  
 aacgtgttaa agaaaaactg caggaaaact acaacaacaa caaaaccatc aaattcgctc 3000  
 cgtcctccgg tgggtgacctg gaaatcacca cccactcctt caactgccgt ggtgaattct 3060  
 tctactgcaa caccgctagc ctgccgtgcc gtatcaaaca gatcatcaac atgtggcagg 3120  
 gtgttggtcg tgctatgtac gctccgccga tcgctggtaa catcacctgc aaatccaaca 3180  
 tcaccggtct gctgctgggt cgtgacggtg gtgaagacaa caaaaccgaa gaaatcttcc 3240  
 gtccgggtgg tggtaacatg aaagacaact ggcgttccga actgtacaaa tacaaagtta 3300  
 tcgaactgaa accgctgggt atcgtccga ccggtgctaa gcttaccgtt caggctcgtc 3360  
 agctgctgtc ctccatcgtt cagcagcagt ccaacctgct gctgctatc gaagctcagc 3420  
 agcacatgct gcagctgacc gtttggggta tcaaacagct gcagaccctg gttctggcta 3480  
 tcgaacgtta cctgaaagac cagcagctcg agatctggaa caacatgacc tggatggaat 3540  
 gggaccgtga aatctccaac tacaccgaca ccatctaccg tctgctggaa gactcccaga 3600  
 cccagcagga aaaaaacgaa aaagacctgc tggctctgga ctctggaaa aacctgtggt 3660  
 cctgggttcga catctccaac tggctgtggt acatcaaaat ctctcatcatg atcgttggtg 3720  
 gtctgatcgg tctgcgtatc atcttcgctg ttctgtccat ctgaggatcc agatctgctg 3780  
 tgccttctag ttgccagcca tctgtgtttt gcccctcccc cgtgccttcc ttgaccctgg 3840  
 aaggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg cattgtctga 3900  
 gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 3960  
 aagacaatag caggcatgct ggggatgcgg tgggctctat gggtaaccag gtgctgaaga 4020  
 attgaccggg ttctcctgg gccagaaaga agcaggcaca tccccttctc tgtgacacac 4080  
 cctgtccacg cccctgggtc ttaggtccag cccactcat aggacactca tagctcagga 4140  
 gggctccgcc ttcaatccca cccgctaaag tacttgagc ggtctctccc tccctcatca 4200  
 gccaccaaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt 4260  
 aagtgcagag ggagagaaaa tgctccaac atgtgaggaa gtaatgagag aaatcataga 4320  
 attttaaggc catgatttaa ggccatcatg gccttaatct tccgcttctc cgctcactga 4380  
 ctgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4440  
 acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca 4500  
 aaaggccagg aaccgtaaaa aggcgcggtt gctggcgttt tccataggc tccgcccccc 4560  
 tgacgagcat cacaaaaatc gacgtcaag tcagaggtgg cgaaaccgca caggactata 4620  
 aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc 4680  
 gcttaccgga tacctgtccg ctttctccc ttccgggaagc gtggcgcttt ctcatagctc 4740  
 acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga 4800  
 acccccgtt cagcccagcc gctgcgcctt atccggtaac tctcgtcttg agtccaacc 4860  
 ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4920  
 gtatgtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct acactagaag 4980  
 aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gagttggtag 5040

```

ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtggt ttttttggtt gcaagcagca 5100
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cgggggtctga 5160
cgctcagtggt aacgaaaact cacgttaagg gatttttggtc atgagattat caaaaaggat 5220
cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa gtatatatga 5280
gtaaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5340
tctatttcgt tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5400
tgaagaaggt gttgctgact cataccagge ctgaatcgcc ccatcatcca gccagaaagt 5460
gaggagacca cggttgatga gagctttgtt gtaggtggac cagttggtga ttttgaactt 5520
ttgctttgcc acggaacggt ctgcgttgct gggaagatgc gtgatctgat ccttcaactc 5580
agcaaaagtt cgatttatc aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc 5640
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5700
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5760
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg 5820
attccgactc gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta 5880
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5940
atcttcttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 6000
tcaaccaaac cgttatcat tctgtattgc gcctgagcga gacgaaatac gcgatcgctg 6060
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6120
tcaacaatat tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg 6180
gggatcgtag tggtagtaaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 6240
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6300
gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 6360
cgatagattg tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa 6420
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 6480
ctcataaacac cccttgatg actgtttatg taagcagaca gttttattgt tcatgatgat 6540
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6600
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6660
tgtattttaga aaaataaaca aataggggtt ccgcgcacat tccccgaaa agtgccacct 6720
gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6780
ccctttcgtc                                     6790

```

&lt;210&gt; 153

&lt;211&gt; 6637

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 153

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tabtaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780

```



cgtcaatggg	agtttgTTTT	ggcaccAAAA	tcaacggggac	tttccAAAAat	gtcgtAACAA	840
ctccgcccc	ttgacgcaaa	tgggCGGTAG	gcgtgtacgg	tgggaggTct	atataagcag	900
agctcgTTta	gtgaaccgTc	agatCGcctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtGacg	taagTaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcTtg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggTg	tgggttattg	accattattg	accactcccc	1200
tattggTgac	gatactttcc	attactaatc	cataacatgg	ctcttttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggTc	ccattttatta	tttacaAAat	cacatataca	acaacgccgt	cccccgTgcc	1380
cgcagtTTTT	attaaacata	gcgtgggGac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctggTc	ccatgcctcc	1500
agcggctcat	ggTcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgCGgtgc	tgttaacggT	ggagggcagT	gtagtctgag	cagtactcgt	tgctgcccg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcaac	gtcgtcgaca	cgtgtgatea	gatatcgcg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tcbtggTatc	ccgcgtAAct	ggccgcagTg	gtggatgtgg	ggTatcctgg	1980
gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggT	ttggaccgac	gctaaaacca	ccctgttctg	cgcttccgac	accaaagcct	2100
acgaccgtga	agttcacAAC	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggAAat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagccgtaac	tgctccttca	2340
acaccaccac	cgAAatccgt	gacaaaaaac	agcagggtta	cgctctgttc	taccgtccgg	2400
acatcgTtct	gctgaaagaa	aaccgtAACa	actccaacaa	ctccgaatac	atcctgatca	2460
actgcaacgc	ttccaccatc	accaggctt	gcccgaagT	taacttcgac	ccgatcccga	2520
tccactactg	cgtcccggt	ggttacgcta	tcctgaaatg	caacaacaaa	accttctccg	2580
gtaaaggTcc	gtgcaacaac	gtttccaccg	ttcagtgcac	ccatggTatc	aaaccggTtg	2640
tttccaccca	gctgtctgtg	aacggTtccc	tggctgaaaa	agaaatcatc	atccgttccg	2700
aaaacctgac	cgacaacgTt	aaaaccatca	tcgttcacct	gaacaaatcc	gttgaaatcg	2760
tttgacccc	tgctagcgt	tactgcaaca	tctccggTtc	caaatggaac	gaaaccctga	2820
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	2880
cgtcctccgg	tggTgacctg	gaaatcacca	cccactcctt	caactgccgt	ggTgaattct	2940
tctactgcaa	caccaccgT	ctgttcaaca	acaacgctac	cgaagacgaa	accatcacc	3000
tgccgtgccg	tatcaaacag	atcatcaaca	tgtggcagg	tgTtggTcgt	gctatgtacg	3060
ctccgccgat	cgtggTaac	atcacctgca	aatccaacat	caccggTctg	ctgctggTtc	3120
gtgacggTgg	tgaagacaac	aaaaccgaag	aatcttccg	tccgggtggT	ggTaacatga	3180
aagacaactg	gcgttccgaa	ctgtacaaat	acaaagtTat	cgaactgaaa	ccgtgggtga	3240
tcgtccgac	cggTgctaag	cttaccgttc	aggtcgtca	gctgctgtcc	tccatcgTtc	3300
agcagcagTc	caacctgctg	cgtgctatcg	aagctcagca	gcacatgctg	cagctgaccg	3360
tttggggTat	caaacagctg	cagaccgTg	ttctggctat	cgaacgttac	ctgaaagacc	3420
agcagctcga	gatctggaac	aacatgacct	ggatggaatg	ggaccgtgaa	atctccaact	3480
acaccgacac	catctaccgt	ctgctggaag	actcccagac	ccagcaggaa	aaaaacgaaa	3540
aagacctgct	ggctctggac	tcctggaaaa	acctgtggTc	ctggTtcgac	atctccaact	3600
ggctgtggTg	aggatccaga	tctgctgtgc	cttctagtTg	ccagccatct	gttgtttgcc	3660
cctcccccgT	gccttccttg	accctggaag	gtgccactcc	cactgtcctt	tcctaataaa	3720
atgaggAAat	tgcatacgcat	tgtctgagTa	ggTgtcatTc	tattctgggg	ggTggggTgg	3780
ggcaggacag	caagggggag	gattgggaag	acaatagcag	gcatagctgg	gatgcggTgg	3840
gctctatggg	taccaggtg	ctgaagaatt	gacccggTtc	ctcctggggc	agaaagaagc	3900
aggcacatcc	ccttctctgt	gacacacct	gtccacgccc	ctggTtctta	gttccagccc	3960
cactcatagg	acactcatag	ctcaggagg	ctccgccttc	aatcccaccc	gctaaagtac	4020
ttggagcggt	ctctccctcc	ctcatcagcc	caccaaacca	aacctagcct	ccaagagtgg	4080

gaagaaatta aagcaagata ggctattaag tgcagagggga gagaaaatgc ctccaacatg 4140  
tgaggaagta atgagagaaa tcatagaatt ttaaggccat gatttaaggc catcatggcc 4200  
ttaatcttcc gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 4260  
ggatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 4320  
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 4380  
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 4440  
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 4500  
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 4560  
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 4620  
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata 4680  
cggtaactat cgtcttgagt ccaacccggg aagacacgac ttatcgccac tggcagcagc 4740  
cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg 4800  
gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 4860  
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 4920  
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga 4980  
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 5040  
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 5100  
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 5160  
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 5220  
gggggggggg cgctgaggtc tgcctcgtga agaagggtgt gctgactcat accaggcctg 5280  
aatcgcccca tcatccagcc agaaagttag ggagccacgg ttgatgagag ctttgttgta 5340  
ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg cgttgteggg 5400  
aagatgcgtg atctgatcct tcaactcagc aaaagtctga tttattcaac aaagccgccg 5460  
tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa ttctgattag 5520  
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 5580  
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 5640  
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 5700  
aatttccccct cgtcaaaaat aagggtatca agtgagaaat caccatgagt gacgactgaa 5760  
tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 5820  
ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg tgattgcgcc 5880  
tgagcgagac gaaatacgcg atcgtgttga aaaggacaat taaaaacagg aatcgaatgc 5940  
aaccggcgca ggaacactgc cagcgcatca acaatatctt cacctgaatc aggatattct 6000  
tctaatacct ggaatgctgt tttcccgggg atcgcagtggt tgagtaacca tgcacatca 6060  
ggagtacgga taaaatgctt gatggtcggg agaggcataa attccgtcag ccagtttagt 6120  
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 6180  
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 6240  
tcgagagccc atttataccc atataaatca gcatccatgt tggaatttaa tcgcggcctc 6300  
gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact gtttatgtaa 6360  
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 6420  
ttttgagaca caacgtggct tttccccccc cccattatt gaagcattta tcagggttat 6480  
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 6540  
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 6600  
acctataaaa ataggcgtat caccaggccc tttcgtc 6637

&lt;210&gt; 154

&lt;211&gt; 6670

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 154

tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	gagacgggtca	60
cagcttgtct	gtaagcggat	gccggggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcggttg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
cccgcctggc	tgaccgcccc	acgacccccg	cccatgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcacgcg	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatgga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctggt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgcg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattgggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatggggtc	ccattttatta	tttacaaatt	cacatatata	acaacgcctg	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtgggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccc	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcgggtgc	tgttaacggc	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	cgtgtgatca	gatatcgccg	ccgctctaga	gatatcgcca	1920
ccatgcgtgt	tcbtggatc	ccgcgtaact	ggccgcagtg	gtggatgtgg	ggtatcctgg	1980
gtttctggat	gatcatcatc	tgccgtgttg	ttggtaacat	gtgggttacc	gtttactacg	2040
gtgttccggt	ttggaccgac	gctaaaacca	ccctgttctg	cgttcccgac	accaaagcct	2100
acgaccgtga	agttcacaa	gtttgggcta	cccacgcttg	cgttccgacc	gacccgaacc	2160
cgcaggaaat	cgttctggaa	aacgttaccg	aaaacttcaa	catgtggaaa	aacgacatgg	2220
ttgaccagat	gcacgaagac	atcatctccc	tgtgggacca	gtccctgaaa	ccgtgcgtta	2280
aactgacccc	gctgtgcgtt	accctgcact	gcaccaacgc	tagccgtaac	tgctccttca	2340
acaccaccac	cgaaatccgt	gacaaaaaac	agcagggtta	cgtctgttct	taccgtccgg	2400
acatcgttct	gctgaaagaa	aaccgtaaca	actccaacaa	ctccgaatac	atcctgatca	2460
actgcaacgc	ttccaccatc	accaggtctt	gcccgaaggt	taacttcgac	ccgatcccca	2520
tccactactg	cgtccggctt	ggttacgcta	tcctgaaatg	caacaacaaa	accttctccg	2580
gtaaagggtcc	gtgcaacaac	gtttccaccg	ttcagtgcac	ccatggatatc	aaaccgggtg	2640
tttccaccca	gctgctgctg	aacggttccc	tggctgaaaa	agaaatcatc	atccgttccg	2700
aaaacctgac	cgacaacgtt	aaaaccatca	tcgttcacct	gaacaaatcc	gttgaaatcg	2760
tttgcacccg	tgctagcgct	tactgcaaca	tctccggttc	caaatggaac	gaaaccctga	2820
aacgtgttaa	agaaaaactg	caggaaaact	acaacaacaa	caaaaccatc	aaattcgctc	2880
cgtcctccgg	tgggtgacctg	gaaatcacca	cccactcctt	caactgccgt	ggatgaattct	2940
tctactgcaa	caccgctagc	ctgccgtgcc	gtatcaacaa	gatcatcaac	atgtggcagg	3000
gtgttggtcg	tgctatgtac	gctccgcgca	tcgctggtaa	catcacctgc	aaatccaaca	3060
tcaccgggtct	gctgctgggt	cgtgacgggtg	gtgaagacaa	caaaaccgaa	gaaatccttc	3120
gtccgggtgg	tggtaacatg	aaagacaact	ggcgttccga	actgtacaaa	tacaaagtta	3180
tcgaactgaa	accgctgggt	atcgctccga	ccggtgctaa	gcttaccgtt	caggctcgctc	3240
agctgctgtc	ctccatcggt	cagcagcagt	ccaacctgct	gcgtgctatc	gaagctcagc	3300



```

agcacatgct gcagctgacc gtttggggta tcaaacagct gcagacccgt gttctggcta 3360
tcgaacgtta cctgaaagac cagcagctcg agatctggaa caacatgacc tggatggaat 3420
gggaccgtga aatctccaac tacaccgaca ccatctaccg tctgctggaa gactcccaga 3480
cccagcagga aaaaaacgaa aaagacctgc tggctctgga ctcttgaaa aacctgtggt 3540
cctgggttcga catctccaac tggctgtggt acatcaaaat ctcatcatg atcgttggtg 3600
gtctgatcgg tctgcgtatc atcttcgctg ttctgtccat ctgaggatcc agatctgctg 3660
tgcttcttag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgacctggt 3720
aaggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg cattgtctga 3780
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 3840
aagacaatag caggcatgct ggggatgcgg tgggctctat gggtagccag gtgctgaaga 3900
attgacccgg ttctcctggt gccagaaaga agcaggcaca tccccttctc tgtgacacac 3960
cctgtccacg cccctgggttc ttagttccag cccactcat aggacactca tagctcagga 4020
gggtctcgcc ttcaatccca ccgctaaag tacttgagc ggtctctccc tccctcatca 4080
gcccacaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag ataggctatt 4140
aagtgcagag ggagagaaaa tgctccaac atgtgaggaa gtaatgagag aaatcataga 4200
attttaagge catgatttaa ggccatcatg gccttaatct tccgcttctc cgctcactga 4260
ctcgtgcgc tgggtcggtc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 4320
acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca 4380
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccatagge tccgcccccc 4440
tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccgga caggactata 4500
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc 4560
gcttaccgga tactgtccg ctttctccc ttcggaagc gtggcgcttt ctcatagctc 4620
acgctgtagg tatctcagtt cgggtgaggt cgttcgctcc aagctgggct gtgtgcacga 4680
acccccggt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaacc 4740
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 4800
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 4860
aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 4920
ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgg ttttttgttt gcaagcagca 4980
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 5040
cgctcagtg aacgaaaact caggttaagg gattttgggtc atgagattat caaaaaggat 5100
cttcacctag atcctttttaa attaaaaatg aagttttaa tcaatctaaa gtatatatga 5160
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 5220
tctatttctg tcatccatag ttgcctgact cggggggggg gggcgctgag gtctgcctcg 5280
tgaagaaggt gttgctgact cataccagc ctgaatcgcc ccatcatcca gccagaaagt 5340
gaggagacca cggttgatga gagctttggt gtaggtggac cagttggtga ttttgaactt 5400
ttgctttgcc acggaacggt ctgcgttgct gggaagatgc gtgatctgat ccttcaactc 5460
agcaaaagtt cgatttatct aacaaagccg ccgtcccgtc aagtcagcgt aatgctctgc 5520
cagtgttaca accaattaac caattctgat tagaaaaact catcgagcat caaatgaaac 5580
tgcaatttat tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat 5640
gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta tcgggtctgcg 5700
attccgactc gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataaggtta 5760
tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa aagcttatgc 5820
atctctttcc agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca 5880
tcaaccaaac cgttatctat tcgtgattgc gcctgagcga gacgaaatac gcgacgctg 5940
ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca 6000
tcaacaatat ttccacctga atcaggatat tcttctaata cctggaatgc tgttttccc 6060
gggatcgagc tgggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggctc 6120
ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 6180
gcaacgctac ctttgccatg ttccagaaac aactctggcg catcgggctt cccatacaat 6240
cgatagattg tcgcacctga ttgcccagca ttatcgcgag cccatttata cccatataaa 6300
tcagcatcca tgttggaatt taatcgcggc ctcgagcaag acgtttccc 6360
ctcataacac ccttgtatt actgtttatg taagcagaca gttttattgt tcatgatgat 6420
atatttttat cttgtgcaat gtaacatcag agattttgag acacaacgtg gctttcccc 6480
ccccccatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa 6540
tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct 6600

```

gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg 6660  
ccctttcgtc 6670

<210> 155

<211> 6547

<212> DNA

<213> Artificial Sequence

<220>

<223> plasmid pVR1012x/s containing HIV genes

<400> 155

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gtccgcggtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctggt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tcccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca ccccttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tttttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttcgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920
ccatgcgtgt tcgtgggtat ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040
gtgttcgggt ttggaccgac gctaaaacca cctgttctg cgcttccgac accaaagcct 2100
acgaccgtga agttcacac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160
cgcaggaaat cgttctggaa aacgttacgc aaaacttcaa catgtggaaa aacgacatgg 2220
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280
aactgacccc gctgtgcgtt accctgcact gcaccaacgc tagcatcaac tgcaacgctt 2340
ccaccatcac ccaggcttgc ccgaaagtta acttcgaccc gatcccgatc cactactgcg 2400
ctccggctgg ttacgctatc ctgaaatgca acaacaaaac cttctccggt aaagggtccg 2460

```



```

gcaacaacgt ttccaccgtt cagtgcaccc atggtatcaa accggttggt tccaccacgc 2520
tgctgctgaa cggttccctg gctgaaaaag aaatcatcat ccgttccgaa aacctgaccg 2580
acaacgttaa aaccatcatc gttcacctga acaaatccgt tgaaatcggt tgcacccgtc 2640
cgaacaacaa cacccgtaaa tccatgcgta tcggtccggg tcagaccttc tacgctaccg 2700
gtgacatcat cggtgacatc cgtcaggcct actgcaacat ctccggttcc aaatggaacg 2760
aaaccctgaa acgtgtttaa gaaaaactgc aggaaaacta caacaacaac aaaaccatca 2820
aattcgctcc gtcctccggt ggtgacctgg aaatcaccac ccactccttc aactgccgtg 2880
gtgaattctt ctactgcaac accgctagcc tgccgtgccg tatcaaacag atcatcaaca 2940
tgtggcaggg tggttggtcg gctatgtacg ctccgccgat cgctggtaac atcacctgca 3000
aatccaacat caccggtctg ctgctgggtc gtgacgggtg tgaagacaac aaaaccgaag 3060
aaatcttccg tccgggtggt ggtaacatga aagacaactg gcgttccgaa ctgtacaaat 3120
acaaagttaa cgaactgaaa ccgctgggta tcgctccgac cgggtgctaag cttaccgttc 3180
aggctcgta gctgctgtcc tccatcgctc agcagcagtc caacctgctg cgtgctatcg 3240
aagctcagca gcacatgctg cagctgaccg tttgggggat caaacagctg cagaccctgt 3300
ttctggctat cgaacgttac ctgaaagacc agcagctcga gatctggaac aacatgacct 3360
ggatggaatg ggaccgtgaa atctccaact acaccgacac catctaccgt ctgctggaag 3420
actcccagac ccagcaggaa aaaaacgaaa aagacctgct ggctctggac tcctggaaaa 3480
acctgtggtc ctggttcgac atctccaact ggctgtgggt aggatccaga tctgctgtgc 3540
cttctagttg ccagccatct gttgtttgcc cctccccctg gccttccttg acctggaag 3600
gtgccactcc cactgtcctt tctaataaaa atgaggaaat tgcctcgcct tgtctgagta 3660
ggtgtcattc tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag 3720
acaatagcag gcatgctggg gatgcgggtg gctctatggg taccaggtg ctgaagaatt 3780
gacccggttc ctctggggcc agaaagaagc aggcacatcc ccttctctgt gacacacct 3840
gtccacgccc ctggttctta gttccagccc cactcatagg acactcatag ctgaggagg 3900
ctccgccttc aatcccaccc gctaaagtac ttggagcggc ctctccctcc ctcatcagcc 3960
caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata ggctattaag 4020
tgagagggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa tcatagaatt 4080
ttaaggccat gatttaaggc catcatggcc ttaatcttcc gcttctctgc tcaactgactc 4140
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4200
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4260
ggccaggaac cgtaaaaagg ccgctgtgct ggcttttttc cataggctcc gccccctga 4320
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgacag gactataaag 4380
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga cctgcccgt 4440
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgcttctc atagctcacg 4500
ctgtaggtat ctgagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4560
ccccgttcag cccgaccgct gcgccttctc cggtaactat cgtcttgagt ccaaccgggt 4620
aagacacgac ttatcgccac tggcagcagc cactggtaac aggatagca gagcgaggta 4680
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4740
agtatttggg atctgcgctc tgetgaagcc agttaccttc ggaaaaagag ttggtagctc 4800
ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat 4860
tacgcgcaga aaaaaaggat ctcaagaaga tcccttgatc ttttctacgg ggtctgacgc 4920
tcagtgaac gaaaactcac gttaagggat tttgggtcat agattatcaa aaaggatctt 4980
cacctagatc cttttaaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5040
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5100
atttcgttca tccatagttg cctgactcgg gggggggggg cgctgaggtc tgcctcgtga 5160
agaagggtgt gctgactcat accaggcctg aatcgcccca tcatccagcc agaaagtgag 5220
ggagccacgg ttgatgagag ctltgttgta ggtggaccag ttggtgattt tgaacttttg 5280
ctttgccacg gaacggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc 5340
aaaagttcga tttattcaac aaagccgccg tcccgtcaag tcagcgtaat gctctgccag 5400
tgttacaacc aattaacca ttctgattag aaaaactcat cgagcatcaa atgaaactgc 5460
aatttattca tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa 5520
ggagaaaact caccgaggca gttccatagg atggcaagat cctgggtatc gtctgcgatt 5580
ccgactcgtc caacatcaat acaacctatt aatttcccct cgtcaaaaat aagggttatca 5640
agtgagaaat caccatgagt gacgactgaa tccggtgaga atggcaaaag cttatgcatt 5700
tctttccaga cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcatca 5760

```

```

accaaaccgt tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta 5820
aaaggacaat tacaacacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca 5880
acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg 5940
atcgcagtgg tgagtaacca tgcacatca ggagtacgga taaaatgctt gatggtcgga 6000
agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca 6060
acgctacctt tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga 6120
tagattgtcg cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca 6180
gcatccatgt tgggaatttaa tcgcggcctc gagcaagacg tttcccgctt aatatggctc 6240
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata 6300
tttttatctt gtgcaatgta acatcagaga ttttgagaca caacgtggct tcccccccc 6360
ccccattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 6420
atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 6480
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgat cagcaggccc 6540
tttcgtc 6547

```

&lt;210&gt; 156

&lt;211&gt; 6679

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 156

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagaactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgcctt ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccattgcctcc 1500
agcggctcat ggtcgcctcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctcttct 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg ccgctctaga gatatcgcca 1920  
ccatgcgtgt tcgtgggtatc ccgcgtaact ggccgcagtg gtggatgtgg ggtatcctgg 1980  
gtttctggat gatcatcatc tgccgtgttg ttggtaacat gtgggttacc gtttactacg 2040  
gtgttccggt ttggaccgac gctaaaacca ccctgttctg cgcttccgac accaaagcct 2100  
acgaccgtga agttcacaaac gtttgggcta cccacgcttg cgttccgacc gaccggaacc 2160  
cgcaggaaat cgttctggaa aacggttaccg aaaacttcaa catgtggaaa aacgacatgg 2220  
ttgaccagat gcacgaagac atcatctccc tgtgggacca gtccctgaaa ccgtgcgtta 2280  
aactgacccc gctgtgcggt accctgcact gcaccaacgc tagccgtaac tgctccttca 2340  
acaccaccac cgaaatccgt gacaaaaaac agcagggtta cgctctgttc taccgtccgg 2400  
acatcgttct gctgaaagaa aaccgtaaca actccaacaa ctccgaatac atcctgatca 2460  
actgcaacgc ttccaccatc acccaggctt gcccgaaggt taacttcgac ccgatcccga 2520  
tccactactg cgctccggct gggtacgcta tcctgaaatg caacaacaaa accttctccg 2580  
gtaaagggtc gtgcaacaac gtttccaccg ttcagtgcac ccatgggtatc aaaccggttg 2640  
tttccaccca gctgctgctg aacggttccc tggtgaaaa agaaatcatc atccgttccg 2700  
aaaacctgac cgacaacggt aaaacctca tctgtcacct gaacaaatcc gttgaaatcg 2760  
tttgcacccg tccgaacaac aacacccgta aatccatgcg tatcggtccg ggtcagacct 2820  
tctacgctac cggtgacatc atcggtgaca tccgtcaggc ttactgcaac atctccggtt 2880  
ccaaatggaa cgaaaccctg aaacgtgtta aagaaaaact gcaggaaaac tacaacaaca 2940  
acaaaacccat caaatctcgt ccgtcctccg gtggtgacct ggaaatcacc acccactcct 3000  
tcaactgccg tgggtgaattc ttctactgca acaccgctag cctgccgtgc cgtatcaaac 3060  
agatcatcaa catgtggcag ggtgttggtc gtgctatgta cgctccgccg atcgctggta 3120  
acatcacctg caaatccaac atcacccgtc tgctgctggt tcgtgacggt ggtgaagaca 3180  
acaaaaccga agaaatcttc cgtccgggtg gtggtaacat gaaagacaac tggcgttccg 3240  
aactgtacaa atacaaagtt atcgaactga aaccgctggg tatcgctccg accggtgcta 3300  
agcttaccgt tcaggctcgt cagctgctgt cctccatcgt tcagcagcag tccaacctgc 3360  
tgcgtgctat cgaagctcag cagcacatgc tgcagctgac cgtttgggggt atcaaacagc 3420  
tgcagacccg tgttctggct atcgaacggt acctgaaaga ccagcagctc gagatctgga 3480  
acaacatgac ctggatggaa tgggaccgtg aaatctccaa ctacaccgac accatctacc 3540  
gtctgctgga agactcccag acccagcagg aaaaaaacga aaaagacctg ctggctctgg 3600  
actcctggaa aaacctgtgg tcctgggtcg acatctccaa ctggctgtgg tgaggatcca 3660  
gatctgctgt gccttctagt tgccagccat ctggtgtttg cccctcccc gtgccttctc 3720  
tgacctgga aggtgccact cccactgtcc ttctctaata aaatgaggaa attgcatcgc 3780  
attgtctgag taggtgtcat tctattctgg ggggtgggggt ggggcaggac agcaaggggg 3840  
aggattggga agacaatagc aggcattgct gggatgcggt gggctctatg ggtacccagg 3900  
tgctgaagaa ttgacccggt tcctcctggg ccagaaagaa gcaggcacat ccccttctct 3960  
gtgacacacc ctgtccacgc ccctgggtct tagttccagc cccactcata ggacactcat 4020  
agctcaggag ggctccgcct tcaatcccac ccgctaaagt acttggagcg gtctctccct 4080  
ccctcatcag cccaccaaac caaacctagc ctccaagagt gggaaagaaat taaagcaaga 4140  
taggtatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 4200  
aatcatagaa ttttaaggcc atgatttaag gccatcatgg ccttaatctt ccgcttctc 4260  
gctcactgac tcgtgctgct cggtcgttcg gctgcggcga gcggtatcag ctactcaaa 4320  
ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 4380  
aggccagcaa aaggccagga accgtaaaaa ggccgcggtt ctggcgtttt tccataggct 4440  
ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 4500  
aggactataa agataccagg cgtttcccc tggaagctcc ctctgctgct ctctgttcc 4560  
gacctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttct 4620  
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg 4680  
tgtgcacgaa ccccccgttc agcccgaccg ctgctcctta tccgtaact atcgtcttga 4740  
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggtt acaggattag 4800  
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta 4860  
cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag 4920  
agttggtagc tcttgatccg gcaaaaaaac caccgctggt agcgggtggt tttttgtttg 4980  
caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatccttga tctttctac 5040

```

gggggtctgac gctcagtgga acgaaaactc acgttaaggg attttgggtca tgagattatc 5100
aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agtttttaa atctctaaag 5160
tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc 5220
agcgatctgt ctatttcgtt catccatagt tgcctgactc gggggggggg ggcgctgagg 5280
tctgcctcgt gaagaagggt ttgctgactc ataccaggcc tgaatcgccc catcatccag 5340
ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat 5400
tttgaacttt tgctttgcca cggaacgggtc tgcgttgctg ggaagatgcg tgatctgatc 5460
cttcaactca gcaaaagttc gattttattca acaaagccgc cgtcccgtca agtcagcgta 5520
atgctctgcc agtggttaca ccaattaacc aattctgatt agaaaaactc atcgagcatc 5580
aaatgaaact gcaattttatt catatcagga ttatcaatac catatttttg aaaaagccgt 5640
ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag atcctggtat 5700
cgggtctgca ttccgactcg tccaacatca atacaacctt ttaattttcc ctcgtcaaaa 5760
ataagggttat caagtgagaa atcaccatga gtgacgactg aatccgggtg gaatggcaaa 5820
agcttatgca tttctttcca gacttggttc acaggccagc cattacgctc gtcataaaaa 5880
tcactcgcat caaccaaac gttattcatt cgtgattgct cctgagcgag acgaaatacg 5940
cgatcgctgt taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact 6000
gccagcgcat caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct 6060
gttttcccg ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaaatgc 6120
ttgatggctg gaagaggcat aaattccgtc agccagttta gtctgaccat ctcactctga 6180
acatcattgg caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc 6240
ccatacaatc gatagattgt cgcacctgat tgcgcgacat tatcgcgagc ccatttatac 6300
ccatataaat cagcatccat gttggaattt aatcgcggcc tcgagcaaga cgtttcccg 6360
tgaatatggc tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattggt 6420
catgatgata tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg 6480
ctttccccc cccccatta ttgaagcatt tatcagggtt attgtctcat gagcggtac 6540
atatttgaat gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa 6600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 6660
atcacgaggc cctttcgtc 6679

```

&lt;210&gt; 157

&lt;211&gt; 6438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 157

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgctg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttgggtcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900

```



agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctggt ttgacctcca 960  
tagaagacac cgggaccgat ccagccctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380  
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctggte ccatgcctcc 1500  
agcggctcat ggctcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gttgccgtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcccttccat gggctctttc 1860  
tgcagtcacc gtctcgcaca tgggcgccccg cgccagcgtg ctgagcggcg gcgagctgga 1920  
ccgctgggag aagatccgcc tgcgccccgg cggcaagaag aagtacaagc tgaagcacat 1980  
cgtgtggggc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040  
cgagggctgc cgccagatcc tgggccagct gcagcccagc ctgcagaccg gcagcgagga 2100  
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcacgcagat 2160  
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaaga 2220  
ggcccagcag gccgcgcgcg acaccggcca cagcaaccag gtgagccaga actaccccat 2280  
cgtgcagaac atccagggcc agatgggtgca ccaggccatc agcccccgca cctgaacgc 2340  
ctgggtgaag gtgggtggag agaaggcctt cagccccgag gtgatccccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgcgc agtgggaccg 2520  
cgtgcacccc gtgcacgcgc gcccctcgc ccccgccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaacc 2640  
ccccatcccc gtgggcgaga tctacaagcg ctggatcctc ctgggcctga acaagatcgt 2700  
gcgcattgtac agccccacca gcatectgga catccgccag ggccccaaag agcccttccg 2760  
cgactacgtg gaccgcttct acaagaccct gcgcgcgcag caggccagcc aggaggtgaa 2820  
gaactggatg accgagaccc tgcctgggtgca gaacgcacaac cccgactgca agaccatcct 2880  
gaaggccctg ggccccgcgc ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cggccccggc cacaaggccc gcgtgctggc cgaggccatg agccaggtga ccaacagcgc 3000  
caccatcatg atgcagcgcg gcaacttccg caaccagcgc aagatcgtga agtgcctcaa 3060  
ctgcggcaag gagggccaca ccgcccgcga ctgcgcgcgc ccccgcaaga agggctgctg 3120  
gaagtgcggc aaggaggggc accagatgaa ggactgcacc gagcgacagg ctaatttttt 3180  
agggaagatc tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga 3240  
gccaacagcc ccaccagaag agagcttcag gtttggggaa gagacaacaa ctccctctca 3300  
gaagcaggag ccgatagaca aggaactgta tcttttagct tccctcagat cactcttttg 3360  
cagcgacccc tgcgcacaat aaagataggg ggccagctga aggaggccct tctagaccag 3420  
gccctggatc cagatctgct gtgccttcta gttgccagcc atctgttgtt tgccccctcc 3480  
ccgtgccttc cttgacctg gaagggtgcca ctcccactgt cctttcctaa taaaatgagg 3540  
aaattgcacg gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagg 3600  
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgct gtgggctcta 3660  
tgggtaccca ggtgctgaag aattgacccg gttcctcctg ggccagaaag aagcaggcac 3720  
atccccctct ctgtgacaca ccctgtccac gccctgggt cttagttcca gccccactca 3780  
taggacactc atagctcagg agggctccgc cttcaatccc acccgctaaa gtacttggag 3840  
cggctctctc ctccctcatc agcccaccaa accaaacct gctccaaga gtgggaagaa 3900  
attaaagcaa gataggctat taagtgcaga gggagagaaa atgcctccaa catgtgagga 3960  
agtaatgaga gaaatcatag aattttaagg ccatcatggc cttaatcttc cgcttccctc 4020  
ctcactgact cgctgcgctc ggctcgttcg ctgcggcgag cggtatcagc tcaactcaag 4080  
gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa 4140  
ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 4200



```

cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 4260
ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg 4320
accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgtttct 4380
catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 4440
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtacta tcgtcttgag 4500
tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 4560
agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 4620
actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 4680
gttggttagct cttgatccgg caaacaacc accgctggta gcggtggttt ttttgtttgc 4740
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 4800
gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca 4860
aaaaggatct tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt 4920
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtggaggc acctatctca 4980
gcgatctgtc tatttcgttc atccatagtt gcctgactcg gggggggggg gcgctgaggt 5040
ctgcctcgtg aagaagggtg tgctgactca taccaggcct gaatcgcccc atcatccagc 5100
cagaaagtga gggagccacg gttgatgaga gctttgttgt aggtggacca gttggtgatt 5160
ttgaactttt gctttgccac ggaacggctc gcgttgctcg gaagatgcgt gatctgatcc 5220
ttcaactcag caaaagtctg atttattcaa caaagccgcc gtcccgtcaa gtcagcgtaa 5280
tgctctgcca gtgttacaac caattaacca attctgatta gaaaaactca tcgagcatca 5340
aatgaaactg caatttatcc atatcaggat tatcaatacc atatttttga aaaagccgtt 5400
tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga tcctggtatc 5460
ggtctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa 5520
taagggtatc aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa 5580
gcttatgcat ttctttccag acttggtcaa caggccagcc attacgctcg tcatcaaaat 5640
cactcgcate aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc 5700
gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg 5760
ccagcgcate aacaatatct tcacctgaat caggatatcc ttctaatacc tggaatgctg 5820
ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg ataaaatgct 5880
tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa 5940
catcattggc aacgctacct ttgccatggt tcagaaacaa ctctggcgca tcgggcttcc 6000
catacaatcg atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc 6060
catataaate agcatccatg ttggaattta atcgcgccct cgagcaagac gtttcccgtt 6120
gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt tttattgttc 6180
atgatgatat atttttatct tgtgcaatgt aacatcagag attttgagac acaacgtggc 6240
tttccccccc ccccattat tgaagcattt atcagggtta ttgtctcatg agcggataca 6300
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 6360
tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 6420
tcacgaggcc ctttcgtc
6438

```

&lt;210&gt; 158

&lt;211&gt; 6486

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing SIV genes

&lt;400&gt; 158

```

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgc tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactat cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300

```

```

tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattgg 420
cccgctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaag 540
tgcccaactg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgcat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta ttacaaaatt cacatataca acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgetgcccg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcatgcacc gtcgtcgacg ccgccaccat gggcgtgcgc aacagcgtgc tgagcggcaa 1920
gaaggccgac gagctggaga agatccgcct gcgccgaac ggcaagaaga agtacatgct 1980
gaagcacgtg gtgtgggccc ccaacgagct ggaccgcttc ggcctggccg agagcctgct 2040
ggagaacaag gagggctgcc agaagatcct gagcgtgctg gcgccgctgg tggccaccgg 2100
cagcgagaac ctgaagagcc tgtacaacac cgtgtgcgtg atctgggtgca tccacgcgga 2160
ggagaagggt aagcacaccg aggaggccaa gcagatcgtg cagcgcacc tggtggtgga 2220
gaccggcacc accgagacca tggccaagac cagccgccc accgcgcca gcagcgagc 2280
cggcggcaac taccgggtgc agcagatcgg cggcaactac gtgcacctgc cgtgagccc 2340
gcgcacctg aacgcctgg tgaagctgat cgaggagaag aagttcggcg ccgaggtggt 2400
gcccggcttc caggccctga gcgagggctg cacgccctac gacatcaacc agatgctgaa 2460
ctgctggggc gaccaccagg ccgcatgca gatcatccgc gacatcatca acgaggaggc 2520
cgccgactgg gacctgcagc acccgagcc cgcccgagc cagggccagc tgcgcgagcc 2580
cagcggcagc gacatcgccg gcaccaccag cagcgtggac gagcagatcc agtggatgta 2640
ccgccagcag aaccgatcc cgggtgggcaa catctaccgc cgtgggatac agctgggct 2700
gcagaagtgc gtgcgcatgt acaaccgcac caacatcctg gacgtgaagc agggcccgaa 2760
ggagcccttc cagagctacg tggaccgct ctacaagagc ctgcgcgccc agcagaccga 2820
cgccgcccgt aagaactgga tgaccagac cctgctgatc cagaacgcca acccgactg 2880
caagctggtg ctgaaggggc tgggcgtgaa ccgaccctg gaggagatgc tgaccgctg 2940
ccagggcgtg ggcggcccgg gccagaaggc ccgctgatg gccgaggccc tgaaggaggc 3000
cctggcgccc gtgcccattc cgttcgcggc cgccagcag cgcgcccgc gcaagcccat 3060
caagtgctgg aactgcggca aggagggcca cagcgcgcc cagtgcgcg cgccgcgccc 3120
ccagggctgc tggaaagtgc gcaagatgga ccacgtgatg gccaagtgcc cggaccgcca 3180
ggcgggtttt ttaggccttg gtccatgggg aaagaagccg cgcaatttcc cgatggctca 3240
agtgcacagc gggctgatgc caactgctcc gccagaggac ccagctgtgg atctgctaaa 3300
gaactacatg cagttgggca agcagcagag agaaaagcag agagaaagca gagagaagcc 3360
ttacaaggag gtgacagagg atttgctgca cctcaattct ctctttggag gagaccagta 3420
gtcactgcat cgatcgaggg ccagcccgtg gaggtgcttc tagaccaggc cctggatcca 3480
gatctgctgt gccttctagt tgccagccat ctgtgtttt cccctcccc gtgccttct 3540
tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc 3600

```

```

attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg 3660
aggattggga agacaatagc aggcattgctg gggatgcggg gggctctatg ggtacccagg 3720
tgctgaagaa ttgacccggg tcctcctggg ccagaaagaa gcaggcacat ccccttctct 3780
gtgacacacc ctgtccacgc ccctgggttct tagttccagc cccactcata ggacactcat 3840
agctcaggag ggctccgcct tcaatcccac ccgctaaagt acttgagagc gtctctccct 3900
ccctcatcag cccaccaaac caaacctagc ctccaagagt ggaagaaat taaagcaaga 3960
taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 4020
aatcatagaa ttttaaggcc atcatggcct taatcttccg ctctctcgct cactgactcg 4080
ctgcgctcgg tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg 4140
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag 4200
gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac 4260
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 4320
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 4380
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc 4440
tgtaggtatc tcagttcggt gtaggtcggt cgctccaagc tgggctgtgt gcacgaacc 4500
ccggttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta 4560
agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggat 4620
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca 4680
gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 4740
tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt 4800
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 4860
cagtggaaacg aaaactcacg ttaagggtatt ttggtcatga gattatcaaa aaggatcttc 4920
acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa 4980
acttggctctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta 5040
tttcgttcat ccatagttgc ctgactcggg gggggggggc gctgaggctc gcctcgtgaa 5100
gaaggtgttg ctgactcata ccaggcctga atcgcccat catccagcca gaaagtgagg 5160
gagccacggg tgatgagagc tttgttgtag gtggaccagt tgggtgatttt gaacttttgc 5220
tttgccacgg aacgggtctgc gttgtcggga agatgcgtga tctgacctt caactcagca 5280
aaagttcgat ttattcaaca aagccgcgct cccgtcaagt cagcgtaatg ctctgccagt 5340
gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa tgaaactgca 5400
atattattcat atcaggatta tcaataccat atttttgaaa aagccgttct tgtaatgaag 5460
gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc 5520
cgactcgtcc aacatcaata caacctatta atttccctc gtcaaaaata aggttatcaa 5580
gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc ttatgcattt 5640
ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa 5700
ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga tcgctgttaa 5760
aaggacaatt acaaacagga atcgaatgca accggcgagc gaacactgcc agcgcatcaa 5820
caatatttct acctgaatca ggatattctt ctaataacct gaatgctgtt tccccgggga 5880
tcgcagtggg gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa 5940
gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa 6000
cgctaccttt gccatgttct agaaacaact ctggcgcacg gggcttccca tacaatcgat 6060
agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca tataaatcag 6120
catccatgtt ggaatttaat cgcggcctcg agcaagacgt tccccgttga atatggctca 6180
taacaccctt tgtattactg tttatgtaag cagacagttt battgttcat gatgatata 6240
ttttatcttg tgcaatgtaa catcagagat tttagacac aacgtggctt tcccccccc 6300
cccattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta 6360
tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg 6420
tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct 6480
ttcgtc
6486

```

&lt;210&gt; 159

&lt;211&gt; 9189

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 159

```

tcgcgcgcttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt battaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttccctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgcggt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat ggggtcttttc 1860
tgagtcacc gtcgtcgaca tgggcgcccc cgcagcgtg ctgagcggcg gcgagctgga 1920
ccgctgggag aagatccgcc tgcgcccccg cggcaagaag aagtacaagc tgaagcacat 1980
cgtgtgggac agccgcgagc tggagcgctt cgcgtgaac cccggcctgc tggagaccag 2040
cgagggctgc cgcagatcc tgggccagct gcagcccagc ctgcagaccg gcagcgagga 2100
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220
ggcccagcag gccgccgccc acaccggcca cagcaaccag gtgagccaga actaccccat 2280
cgtgcagAAC atccagggcc agatgggtgca ccaggccatc agcccccgca cctgaacgc 2340
ctgggtgaag gtggtggagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400
cctgagcgag ggcgccacc ccaggaacct gaacaccatg ctgaacaccg tgggcggcca 2460
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgccc agtgggaccg 2520
cgtgcacccc gtgcacgcc gccccatcgc ccccgccag atgcgcgagc cccgcggcag 2580
cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaacc 2640
ccccatcccc gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700
gcgcagtac agccccacca gcatcctgga catccgcccag ggccccagg agcccttccg 2760
cgactacgtg gaccgcttct acaagacct gcgcgccgag caggccagcc aggaggtgaa 2820
gaactggatg accgagaccc tgcgtggtgca gaacgccaac cccgactgca agaccatcct 2880
gaaggccctg ggcgccgccc ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940

```



cggccccggc	cacaaggccc	gcgtgctggc	cgaggccatg	agccaggtga	ccaacagcgc	3000
caccatcatg	atgcagcgcg	gcaacttccg	caaccagcgc	aagatcgtga	agtgttcaa	3060
ctgcggcaag	gagggccaca	ccgcccgcaa	ctgccgcgcc	ccccgcaaga	agggctgctg	3120
gaagtgcggc	aaggagggcc	accagatgaa	ggactgcacc	gagcgacagg	ctaataggga	3180
agatctggcc	ttcccacaag	ggaaggccag	ggaattttct	tcagagcaga	ccagagccaa	3240
cagccccacc	agaagagagc	ttcaggtttg	gggaagagac	aacaactccc	tctcagaagc	3300
aggagccgat	agacaaggaa	ctgtatcctt	tagcttccct	cagatcactc	tttggcagcg	3360
accctcgtc	acaataaaga	tagggggcca	gctgaaggag	gcccttctag	acaccggcgc	3420
cgacgacacc	gtgctggagg	agatgaacct	gcccggccgc	tgggaagccca	agatgatcgg	3480
cggcatcggc	ggcttcatca	aggtggggcca	gtacgaccag	atcctgatcg	agatctgcgg	3540
ccacaaggcc	atcggcaccg	tgctgggtggg	ccccaccccc	gtgaacatca	tcggccgcaa	3600
cctgctgacc	cagatcggct	gcaccctgaa	cttccccatc	agccccatcg	agaccgtgcc	3660
cgtgaagctg	aagcccggca	tggacggccc	caaggtgaag	cagtggcccc	tgaccgagga	3720
gaagatcaag	gccctggtgg	agatctgcac	cgagatggag	aaggagggca	agatcagcaa	3780
gatcggcccc	gagaacccct	acaacacccc	cgtgttcgcc	atcaagaaga	aggacagcac	3840
caagtggcgc	aagctggtgg	acttccgcga	gctgaacaag	cgcacccagg	acttctggga	3900
ggtgcagctg	ggcatccccc	accccgccgg	cctgaagcag	aagaagagcg	tgaccgtgct	3960
ggacgtgggc	gacgcctact	tcagcgtgcc	cctggacaag	gacttccgca	agtacaccgc	4020
cttcaccatc	cccagcatca	acaacgagac	ccccggcatc	cgctaccagt	acaacgtgct	4080
gccccagggc	tggaaaggca	gccccgccat	cttccagtgc	agcatgacca	agatcctgga	4140
gcccttccgc	aagcagaacc	ccgacatcgt	gatctaccag	tacatggacg	acctgtacgt	4200
gggcagcgac	ctggagatcg	gccagcaccg	caccaagatc	gaggagctgc	gccagcacct	4260
gctgcgctgg	ggcttcacca	cccccgacaa	gaagcaccag	aaggagcccc	ccttcctgtg	4320
gatgggctac	gagctgcacc	ccgacaagtg	gaccgtgcag	cccatcgtgc	tgcccagaaa	4380
ggacagctgg	accgtgaacg	acatccagaa	gctgggtggc	aagctgaact	gggccagcca	4440
gatctacgcc	ggcatcaagg	tgcgccagct	gtgcaagctg	ctgcgcggca	ccaaggccct	4500
gaccgaggtg	gtgcccctga	ccgaggaggc	cgagctggag	ctggccgaga	accgcgagat	4560
cctgaaggag	cccgtgcacg	gcgtgtacta	cgaccccagc	aaggacctga	tcgcccagat	4620
ccagaagcag	ggccagggcc	agtggacctt	ccagatctac	caggagccct	tcaagaacct	4680
gaagaccggc	aagtacgccc	gcatgaaggg	cgccacacac	aacgacgtga	agcagctgac	4740
cgaggccgtg	cagaagatcg	ccaccgagag	catcgtgatc	tggggcaaga	cccccaagtt	4800
caagctgccc	atccagaagg	agacctggga	ggcctggtgg	accgagtact	ggcaggccac	4860
ctggatcccc	gagtgggagt	tcgtgaacac	cccccccctg	gtgaagctgt	ggtaccagct	4920
ggagaaggag	cccatcatcg	gcgcgcgagc	cttctacgtg	gacggcgccg	ccaaccgcga	4980
gaccaagctg	ggcaaggccg	gctacgtgac	cgaccgcggc	cgccagaagg	tggtgcccct	5040
gaccgacacc	accaaccaga	agaccgagct	gcaggccatc	cacctggccc	tgaggagacg	5100
cggcctggag	gtgaacatcg	tgaccgacag	ccagtacgcc	ctgggcatca	tccaggccca	5160
gccccgacaag	agcgagagcg	agctgggtgag	ccagatcatc	gagcagctga	tcaagaagga	5220
gaagggtgtac	ctggcctggg	tgcccgccca	caagggcatc	ggcggcaacg	agcagggtga	5280
cggcctgggtg	agcgccggca	tccgcaagggt	gctgttcctg	gacggcatcg	acaaggccca	5340
ggaggagcac	gagaagtacc	acagcaactg	gcgcgccatg	gccagcgact	tcaacctgcc	5400
ccccgtgggtg	gccaaggaga	tcgtggccag	ctgcgacaag	tgccagctga	agggcgaggc	5460
catgcacggc	caggtggact	gcagccccgg	catctggcag	ctggactgca	cccacctgga	5520
gggcaagggtg	atcctgggtg	ccgtgcacgt	ggccagcggc	tacatcgagg	ccgagggtgat	5580
ccccgcccag	accggccagg	agaccgccta	cttccctgctg	aagctggccg	gccgctggcc	5640
cgtgaagacc	gtgcacaccg	acaacggcag	caacttcacc	agcaccaccg	tgaaggccgc	5700
ctgctgggtg	gccggcatca	agcaggagtt	cggcatcccc	tacaaccccc	agagccaggg	5760
cgtgatcgag	agcatgaaca	aggagctgaa	gaagatcatc	ggccagggtg	gcgaccaggc	5820
cgagcacctg	aagaccgccc	tgcatatggc	cgtgttcatc	cacaacttca	agcgcaaggg	5880
cggcatcggc	ggctacagcg	ccggcgagcg	catcgtggac	atcatcgcca	ccgacatcca	5940
gaccaaggag	ctgcagaagc	agatcaccaa	gatccagaac	ttccgcgtgt	actaccgcga	6000
cagccgcgac	cccgtgtgga	agggccccgc	caagctgctg	tggaaaggcg	agggcgccgt	6060
ggtgatccag	gacaacagcg	acatcaagggt	ggtgccccgc	cgcaaggcca	agatcatccg	6120
cgactacggc	aagcagatgg	ccggcgacga	ctgcgtggcc	agccgcccag	acgaggacta	6180
ggaattctgc	tgtgccttct	agttgccagc	catctgttgt	ttgcccctcc	cccgtgcctt	6240



```

ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat 6300
cgcatgtct gagtaggtgt cattctatct tggggggtgg ggtggggcag gacagcaagg 6360
gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atgggtaccc 6420
aggtgctgaa gaattgaccc ggttcctcct gggccagaaa gaagcaggca catccccctc 6480
tctgtgacac accctgtcca cgcccctggg tcttagttcc agcccactc ataggacact 6540
catagctcag gagggctccg ccttcaatcc caccgcctaa agtacttgga gcggtctctc 6600
cctccctcat cagcccacca aaccaaacct agcctccaag agtgggaaga aattaaagca 6660
agataggcta ttaagtgcag agggagagaa aatgcctcca acatgtgagg aagtaatgag 6720
agaaatcata gaattttaag gccatcatgg ccttaatctt ccgcttcctc gctcactgac 6780
tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctactcaaaa ggcggtaata 6840
cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 6900
aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccctc 6960
gacgagcatc aaaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa 7020
agataccagg cgtttccccc tggaagctcc ctctgctgct ctctgttcc gaccctgccg 7080
cttaccggat acctgtccgc ctttctcctc tcgggaagcg tggcgctttc tcatagctca 7140
cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 7200
cccccgctt agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg 7260
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 7320
tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaaga 7380
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 7440
tcttgatccg gcaaacaac caccgctggg agcgggtggg tttttgtttg caagcagcag 7500
attacgcgca gaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac 7560
gctcagtggg acgaaaactc acgttaaggg attttgggtc tgagattatc aaaaaggatc 7620
ttcacctaga tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag 7680
taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt 7740
ctatttcgtt catccatagt tgcctgactc gggggggggg ggcgctgagg tctgcctcgt 7800
gaagaaggtg ttgctgactc ataccaggcc tgaatcgccc catcatccag ccagaaagtg 7860
aggagaccac gggtgatgag agctttgttg taggtggacc agttgggtgat tttgaacttt 7920
tgctttgcca cggaacggtc tgcgttgctg ggaagatgcg tgatctgac cttcaactca 7980
gcaaaagttc gatttattca acaaagccgc cgtcccgtca agtcagcgta atgctctgcc 8040
agtggttaca ccaattaacc aattctgatt agaaaaactc atcgagcatc aaatgaaact 8100
gcaatttatt catatcagga ttatcaatac catatttttg aaaaagccgt ttctgtaatg 8160
aaggagaaaa ctcccgagg cagttccata ggatggcaag atcctgggat cggctctgca 8220
ttccgactcg tccaacatca atacaacct ttaattttcc ctctcaaaa ataaggttat 8280
caagtgagaa atcaccatga gtgacgactg aatccgggtg gaatggcaaa agcttatgca 8340
tttctttcca gacttggtca acaggccagc cattacgctc gtcacaaaa tcaactcgc 8400
caaccaaac gttattcatt cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt 8460
taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgcat 8520
caacaatatt ttcacctgaa tcaggatatt cttctaatac ctggaatgct gttttcccg 8580
ggatcgcat ggtgagtaac catgcatcat caggagtacg gataaaatgc ttgatggctg 8640
gaagaggcat aaattccgtc agccagttta gtctgacct ctcatctgta acatcattgg 8700
caacgctacc tttgccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc 8760
gatagattgt cgcacctgat tgcccagac tatcgcgagc ccatttatac ccataaaat 8820
cagcatccat gttggaattt aatcgcgggc tcgagcaaga cgtttcccgt tgaatatggc 8880
tcataacacc ccttgattta ctgtttatgt aagcagacag ttttattggt catgatgata 8940
tatttttatc ttgtgcaatg taacatcaga gattttgaga cacaacgtgg ctttcccccc 9000
cccccattha ttgaagcatt tatcagggtt attgtctcat gagcggtac atatttgaat 9060
gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg 9120
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 9180
cctttcgtc
9189

```

&lt;210&gt; 160

&lt;211&gt; 9285

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing SIV genes

&lt;400&gt; 160

```
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaag 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acgggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttatgt accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatat tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgcctg ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgcg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcatgcacc gtcgtcgacg ccgccaccat gggcgtgcgc aacagcgtgc tgagcggcaa 1920
gaaggccgac gagctggaga agatccgcct gcgcccgaac ggcaagaaga agtacatgct 1980
gaagcacgtg gtgtggggcc ccaacgagct ggaccgcttc ggcttgcccg agagcctgct 2040
ggagaacaag gagggctgcc agaagatcct gagcgtgctg gcgcccgtgg tgcccaccgg 2100
cagcgagaac ctgaagagcc tgtacaacac cgtgtgcgtg atctgggtga tccacgccga 2160
ggagaagggt aagcacaccg aggaggccaa gcagatcgtg cagcgccacc tgggtggtgga 2220
gaccggcacc accgagacca tgcccaagac cagccgcccc acccgcccca gcagcggacg 2280
cggcggcaac taccgggtgc agcagatcgg cggcaactac gtgcacctgc cgtgagccc 2340
gcgcaccctg aacgcctggg tgaagctgat cgaggagaag aagttcggcg ccgaggtggt 2400
gcccggcttc caggccctga gcgagggctg cacgcccctac gacatcaacc agatgctgaa 2460
ctgctggtgg gaccaccagg ccgccatgca gatcatccgc gacatcatca acgaggaggc 2520
cgccgactgg gacctgcagc acccgagacc cgcgcgcgag cagggccagc tgcgcgagcc 2580
cagcggcagc gacatcgccg gcaccaccag cagcgtggac gagcagatcc agtggatgta 2640
ccgccagcag aaccgatcc cgggtgggcaa catctaccgc cgtgggatac agctgggcct 2700
gcagaagtgc gtgcgcatgt acaaccgcag caacatcctg gacgtgaagc agggcccgaa 2760
ggagcccttc cagagctacg tggaccgctt ctacaagagc ctgcgcgccc agcagaccga 2820
cgccgcgctg aagaactgga tgaccagac cctgctgata cagaacgcca acccgactg 2880
```

caagctggtg ctgaagggcc tgggctgaa cccgacctg gaggagatgc tgaccgcctg 2940  
 ccagggcgtg ggcgggcccg gccagaaggc ccgctgatg gccgaggccc tgaaggaggc 3000  
 cctggcgccc gtgcccaccc cgttcgcggc cgcccagcag cgcgggcccg gcaagcccat 3060  
 caagtgctgg aactgcggca aggagggcca cagcgcccgc cagtgcgcgc cgccgcgcgc 3120  
 ccagggctgc tggagtgcg gcaagatgga ccacgtgatg gccagtgcc cggaccgcca 3180  
 ggcgggtttt ttaggccttg gtccatgggg aaagaagccg cgcaatttcc cgatggctca 3240  
 agtgcacag gggctgatgc caactgctcc gccagaggac ccagctgtgg atctgctaaa 3300  
 gaactacatg cagttgggca agcagcagag agaaaagcag agagaaagca gagagaagcc 3360  
 ttacaaggag gtgacagagg atttgctgca cctcaattct ctctttggag gagaccagta 3420  
 gtcactgcat cgatcgaggg ccagcccgtg gaggtgcttc tagacaccgg cgccgacgac 3480  
 agcatcgtga ccggcatcga gctggggccc cactacacc cgaagatcgt gggcggcac 3540  
 ggcggcttca tcaacaccgg agagtacaag aacgtggaga tcgaggtgct gggcaagcgc 3600  
 atcaagggca ccatcatgac cggcgacacg cccatcaaca tcttcggccg caacctgctg 3660  
 accgccctgg gcatgagcct gaacttcccc atcgccaagg tggagcccgt gaaggtggcc 3720  
 ctgaagcccc gcaaggacgg cccaagctg aagcagtgcc cctgagcaa ggagaagatc 3780  
 gtggccctgc gcgagatctg cgagaagatg gagaaggacg gccagctgga ggaggcccca 3840  
 cccaccaacc cctacaacac gccaccttc gccatcaaga agaaggacaa gaacaagtgg 3900  
 cgcattgctga tcgacttccg cgagctgaac cgcgtgaccc aggacttcac cgaggtgcag 3960  
 ctgggcatcc cgcaccccg cggcctggcc aagcgcaagc gcatcacctg gctggacatc 4020  
 ggcgacgcct acttcagcat cccctggac gaggagtcc gccagtacac cgccttcacc 4080  
 ctgcccagcg tgaacaacgc cgagcccggc aagcgctaca tctacaagg tctgccccag 4140  
 ggctggaagg gcagccccgc catcttccag tacaccatgc gccacgtgct ggagcccttc 4200  
 cgcaaggcca accccgacgt gacctggtg cagtacatgc atgacatcct gatcgccagc 4260  
 gaccgcaccg acctggagca cgaccgctg gtgctgcaga gcaaggagct gctgaacagc 4320  
 atcggttca gcagccccga ggagaagttc cagaaggacc cacccttcca gtggatgggc 4380  
 tacgagctgt ggcccaccaa gtggaagctg cagaagatcg agctgcccc a gcgcgagacc 4440  
 tggaccgtga acgacatcca gaagctggtg ggcgtgctga actgggccc cagatctac 4500  
 cccggcatca agaccaagca cctgtgccgc ctgatccgc gcaagatgac cctgaccgag 4560  
 gaggtgcagt ggaccgagat ggccgaggcc gagtacgagg agaacaagat catcctgagc 4620  
 caggagcagg agggctgcta ctaccaggag ggcaagcccc tggaggccac cgtgatcaag 4680  
 agccaggaca accagtggag ctacaagatc caccaggagg acaagatcct gaaggtgggc 4740  
 aagttcgcca agatcaagaa caccacacc aacggcgtgc gcctgctggc ccacgtgatc 4800  
 cagaagatcg gcaaggaggc catcgtgatc tggggccagg tgcccaagtt ccacctgccc 4860  
 gtggagaagg acgtgtggga gcagtgggtg accgactact ggcaggtgac ctggatacc 4920  
 gagtgggact tcatcagcac cctccccctg gtgcgcctgg tgttcaacct ggtgaaggac 4980  
 cccatcgagg gcgaggagac ctactacacc gacggcagct gcaacaagca gagcaaggag 5040  
 ggcaaggccg gctacatcac cgaccgcggc aaggacaagg tgaagggtgct ggagcagacc 5100  
 accaaccagc aggcgagct cgaggccttc ctgatggccc tgaccgacag cggccccaag 5160  
 gccaacatca tcgtggacag ccagtacgtg atgggcatca tcaccggctg cccaccgag 5220  
 agcgagagcc gcctggtgaa ccagatcatc gaggagatga tcaagaagag cgagatctac 5280  
 gtggcctggg tgcccgccea caagggcac tggcggaacc aggagatcga ccacctggtg 5340  
 agccagggca tccgccagg tctgttctct gagaagatcg agcccgccea ggaggagcac 5400  
 gacaagtacc acagcaacgt gaaggagctg gtgttcaagt tcggcctgcc ccgcatcgtg 5460  
 gcccgcaga tcgtggacac ctgcaacaag tgccaccaga agggcgaggc catccacggc 5520  
 caggccaaca gcgacctggg cacctggcag atggcctgca cccacctgga gggcaagatc 5580  
 atcatcgtg ccgtgcacgt ggccagcggc ttcacagagg ccgaggtgat cccgcaggag 5640  
 accggccgcc agaccgccct gtctctgctg aagctggccg gccgctggcc catcaccac 5700  
 ctgcacaccg acaacggcgc caacttcgcc agccaggagg tgaagatggt ggcctggtg 5760  
 gccggcatcg agcacacctt cggcgtgccc tacaaccgc agagccaggg cgtggtggag 5820  
 gccatgaacc accacctgaa gaaccagatc gaccgcaccc gcgagcaggg caacagcgtg 5880  
 gagaccatcg tgctgatggc cgtgcactgc atgaacttca agcgcgcgcg cggcatcggc 5940  
 gacatgacgc ccgccgagcg cctgatcaac atgatacca ccgagcagga gatccagtcc 6000  
 cagcagagca agaacagcaa gttcaagaac ttccgcgtgt actaccgca gggccgcgac 6060  
 cagctgtgga agggccccgg cgagctgctg tggaggggcg agggcgccgt gatcctgaag 6120  
 gtgggcaccg acatcaagg tggccccgc cgcaaggcca agatcatcaa ggactacggc 6180



ggcggcaagg	aggtggacag	cagcagccac	atggaggaca	ccggcgagge	ccgcgaggtg	6240
gcctgagaat	tcgttttaac	acgcgtgtat	acggatccag	atctgctgtg	ccttctagtt	6300
gccagccatc	tgttggttgc	ccctcccccg	tgccttcctt	gacctggaa	ggtgccactc	6360
ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	ttgtctgagt	aggtgtcatt	6420
ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	6480
ggcatgctgg	ggatgcgggtg	ggctctatgg	gtacccaggt	gctgaagaat	tgacccgggt	6540
cctcctgggc	cagaaagaag	caggcacatc	cccttctctg	tgacacacce	tgtccacgcc	6600
cctgggttctt	agttccagcc	ccactcatag	gacactcata	gctcaggagg	gctccgcctt	6660
caatcccacc	cgctaaagta	cttgagcgcg	tctctccctc	cctcatcagc	ccaccaaac	6720
aaacctagcc	tccaagagtg	ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	6780
agagaaaatg	cctccaacat	gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	6840
tcattgacct	aatcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	6900
cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	6960
aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	7020
gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	7080
tcaagtcaga	ggtggcgaaa	cccgcacagga	ctataaagat	accaggcggt	tccccctgga	7140
agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	7200
ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggatatc	cagttcgggtg	7260
taggtcggtc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	7320
gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	7380
gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	taggcgggtg	tacagagttc	7440
ttgaagtggg	ggcctaacta	cggctacact	agaagaacag	tatttggtat	ctgcgctctg	7500
ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	7560
gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	cgcgagaaaa	aaaaggatct	7620
caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaaacg	aaactcacgt	7680
taagggattt	tggatcatgag	attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	7740
aaatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	cttgggtctga	cagttacca	7800
tgcttaatac	gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcatc	catagttgcc	7860
tgactcgggg	ggggggggcg	ctgaggtctg	cctcgtgaag	aagggtgttg	tgactcatat	7920
caggcctgaa	tcgccccatc	atccagccag	aaagtgaggg	agccacgggt	gatgagagct	7980
ttgttgtagg	tggaccagtt	ggtgattttg	aacttttgct	ttgccacgga	acgggtctgcg	8040
ttgtcgggaa	gatgcgtgat	ctgatccttc	aactcagcaa	aagttcgatt	tattcaacaa	8100
agccgcgctc	ccgtcaagtc	agcgtaatgc	tctgccagtg	ttacaaccaa	ttaaccaatt	8160
ctgattagaa	aaactcatcg	agcatcaaat	gaaactgcaa	tttattcata	tcaggattat	8220
caataccata	tttttgaaaa	agccgtttct	gtaatgaagg	agaaaactca	ccgaggcagt	8280
tccataggat	ggcaagatcc	tggtatcggt	ctgcgattcc	gactcgtcca	acatcaatac	8340
aacctattaa	tttccccctg	tcaaaaataa	ggttatcaag	tgagaaatca	ccatgagtga	8400
cgactgaatc	cgggtgagaat	ggcaaaagct	tatgcatttc	tttccagact	tgttcaacag	8460
gccagccatt	acgctcgtca	tcaaaatcac	tcgcatcaac	caaaccgtta	ttcattcgtg	8520
attgcgctcg	agcgagacga	aatacgcgat	cgctgttaaa	aggacaatta	caaacaggaa	8580
tcgaatgcaa	ccggcgaggg	aacactgcca	gcgcatcaac	aatattttca	cctgaatcag	8640
gatattcttc	taatacctgg	aatgctgttt	tcccggggat	cgcagtgggt	agtaaccatg	8700
catcatcagg	agtacggata	aaatgcttga	tggctcggaag	aggcataaat	tccgtcagcc	8760
agtttagtct	gaccatctca	tctgtaacat	cattggcaac	gctacctttg	ccatgtttca	8820
gaaacaactc	tggcgcatcg	ggcttcccat	acaatcgata	gattgtcgca	cctgattgcc	8880
cgacattatc	gcgagcccat	ttatacccat	ataaatcagc	atccatgttg	gaatttaatc	8940
gcggcctcga	gcaagacgtt	tcccgttgaa	tatggctcat	aacacccctt	gtattactgt	9000
ttatgtaagc	agacagtttt	attgttcatg	atgatataat	tttatcttgt	gcaatgtaac	9060
atcagagatt	ttgagacaca	acgtggcttt	ccccccccc	ccattattga	agcatttatc	9120
agggttattg	tctcatgagc	ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	9180
gggttccgcg	cacatttccc	cgaaaagtgc	cacctgacgt	ctaagaaacc	attattatca	9240
tgacattaac	ctataaaaaat	aggcgtatca	cgaggccctt	tcgtc		9285

&lt;210&gt; 161

&lt;211&gt; 7897

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 161

```
tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcggt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg gccattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcgggtg tggttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat ggggtcttttc 1860
tgcatgcacc gtcgtcgaca tgagggaaga tctggccttc ccacaaggga aggccaggga 1920
atthttcttc gagcagacca gagccaacag cccaccaga agagagcttc aggtttgggg 1980
aagagacaac aactccctct cagaagcagg agccgataga caaggaactg tatcctttag 2040
cttccctcag atcactcttt ggcagcgacc cctcgtcaca ataaagatag ggggccagct 2100
gaaggaggcc cttctagaca ccggcgccga cgacaccgtg ctggaggaga tgaacctgcc 2160
cggccgctgg aagcccaaga tgatcggcgg catcggcggc ttcatacaagg tgggccagta 2220
cgaccagatc ctgatcgaga tctgcggcca caaggccatc ggcaccgtgc tgggtggccc 2280
caccctcggtg aacatcatcg gccgcaacct gctgacctag atcggctgca ccctgaactt 2340
ccccatcagc cccatcgaga ccgtgcccgt gaagctgaag cccggcatgg acggcccaaa 2400
ggtgaagcag tggcccttga ccgaggagaa gatcaaggcc ctggtggaga tctgcaccga 2460
gatggagaag gagggcaaga tcagcaagat cggccccgag aacccttaca acaccccgct 2520
gttcgccatc aagaagaagg acagcaccaa gtggcgcaag ctggtggact tccgcgagct 2580
gaacaagcgc acccaggact tctgggaggt gcagctgggc atccccacc ccgcggcct 2640
gaagcagaag aagagcgtga ccgtgctgga cgtggcgac gcctacttca gctgccccct 2700
ggacaaggac ttccgcaagt acaccgcctt caccatcccc agcatcaaca acgagacccc 2760
```



cggcatccgc taccagtaca acgtgctgcc ccagggctgg aagggcagcc ccgccatctt 2820  
 ccagtgcagc atgaccaaga tcctggagcc ctcccgcaag cagaaccccg acatcgtgat 2880  
 ctaccagtac atggacgacc tgtacgtggg cagcgacctg gagatcggcc agcaccgcac 2940  
 caagatcgag gagctgcgcc agcacctgct gcgctggggc ttcaccaccc ccgacaagaa 3000  
 gcaccagaag gagccccctt tcctgtggat gggctacgag ctgcaccccg acaagtggac 3060  
 cgtgcagccc atcgtgctgc ccgagaagga cagctggacc gtgaacgaca tccagaagct 3120  
 ggtgggcaag ctgaactggg ccagccagat ctacgccggc atcaagggtg gccagctgtg 3180  
 caagctgctg cgcggcacca aggccctgac cgaggtggtg cccctgaccg aggaggccga 3240  
 gctggagctg gccgagaacc gcgagatcct gaaggagccc gtgcacggcg tgtactacga 3300  
 ccccagcaag gacctgatcg ccgagatcca gaagcagggc cagggccagt ggacctacca 3360  
 gatctaccag gagcccttca agaacctgaa gaccggcaag tacgcccgca tgaaggcgcg 3420  
 ccacaccaac gacgtgaagc agctgaccga ggccgtgcag aagatcgcca ccgagagcat 3480  
 cgtgatctgg ggcaagaccc ccaagttcaa gctgcccata cagaaggaga cctgggaggc 3540  
 ctggtggacc gagtactggc aggccacctg gatccccgag tgggagttcg tgaacacccc 3600  
 cccctgggtg aagctgtggt accagctgga gaaggagccc atcatcggcg ccgagacctt 3660  
 ctacgtggac ggcgcccgca accgcgagac caagctgggc aaggccggct acgtgaccga 3720  
 ccgcggccgc cagaagggtg tgcccctgac cgacaccacc aaccagaaga ccgagctgca 3780  
 ggccatccac ctggccctgc aggacagcgg cctggagggtg aacatcgtga ccgacagcca 3840  
 gtacgccctg ggcatactcc aggccagacc cgacaagagc gagagcgagc tggtagacca 3900  
 gatcatcgag cagctgatca agaaggagaa ggtgtacctg gcctgggtgc ccgccacaa 3960  
 gggcatcggc ggcaacgagc aggtggacgg cctggtgagc gccggcatcc gcaagggtgt 4020  
 gttcctggac ggcatacgaca aggccagga ggagcacgag aagtaccaca gcaactggcg 4080  
 cgccatggcc agcgacttca acctgcccc cgtggtggcc aaggagatcg tggccagctg 4140  
 cgacaagtgc cagctgaagg gcgaggccat gcacggccag gtggactgca gcccggcat 4200  
 ctggcagctg gactgcaccc acctggaggg caaggtgatc ctggtggccg tgcacgtggc 4260  
 cagcggctac atcgaggccg aggtgatccc cgccgagacc ggccaggaga ccgcctactt 4320  
 cctgctgaag ctggccggcc gctggcccggt gaagaccgtg cacaccgaca acggcagcaa 4380  
 cttcaccagc accaccgtga aggcgcctg ctggtgggccc ggcataagc aggagtccg 4440  
 catcccctac aacccccaga gccagggcgt gatcgagagc atgaacaagg agctgaagaa 4500  
 gatcatcggc caggtgcgcg accaggccga gcacctgaag accgccgtgc agatggccgt 4560  
 gttcatccac aacttcaagc gcaaggggcg catcggcggc tacagcgccg gcgagcgcat 4620  
 cgtggacatc atcgccaccg acatccagac caaggagctg cagaagcaga tcaccaagat 4680  
 ccagaacttc cgcgtgtact accgcgacag ccgcgacccc gtgtggaagg gcccgcgcaa 4740  
 gctgctgtgg aagggcgagg gcgcctgggt gatccaggac aacagcgaca tcaagggtgt 4800  
 gcccgcgcgc aaggccaaga tcatccgcga ctacggcaag cagatggccg gcgacgactg 4860  
 cgtggccagc cgccaggacg aggactagga atttctgctg tgccctctag ttgccagcca 4920  
 tctgttgttt gcccctcccc cgtgccttcc ttgacctgg aaggtgccac tcccactgtc 4980  
 ctttcctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg 5040  
 gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct 5100  
 ggggatgcgg tgggtcttat gggtaaccag gtgctgaaga attgaccgg ttctcctgg 5160  
 gccagaaaga agcaggcaca tccccttctc tgtgacacac cctgtccacg cccctggttc 5220  
 ttagttccag cccactcat aggacactca tagctcagga gggctccgcc ttcaatccca 5280  
 cccgctaaag tacttggagc ggtctctccc tccctcatca gccaccaaaa ccaaacctag 5340  
 cctccaagag tgggaagaaa ttaaagcaag ataggctatt aagtgcagag ggagagaaaa 5400  
 tgccccaac atgtgaggaa gtaatgagag aaatcataga attttaaggc catcatggcc 5460  
 ttaatcttcc gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc 5520  
 ggtatcagct cactcaaagg cggtaatagc gttatccaca gaatcagggg ataacgcagg 5580  
 aaagaacatg tgagcaaaa ggcagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 5640  
 ggcgttttcc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 5700  
 gagggtggcg aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 5760  
 cgtgcgctct cctgttccga cctgcccgt taccggatac ctgtccgcct ttctcccttc 5820  
 ggggaagcgt gcgcttctc atagctcacg ctgtaggtat ctgagttcgg ttaggtcgt 5880  
 tcgctccaag ctgggctgtg tgcacgaacc cccgcttcag ccgaccgct gcgccttctc 5940  
 cggtaactat cgtcttgagt ccaacccggg aagacacgac ttatcgccac tggcagcagc 6000  
 cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 6060

```

gtggcctaac tacggctaca ctagaagaac agtatttggg atctgcgctc tgctgaagcc 6120
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 6180
cgggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 6240
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 6300
tttggtcatt agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag 6360
ttttaaatca atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat 6420
cagtgaaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactcgg 6480
gggggggggg cgctgaggtc tgccctcgtga agaagggtgt gctgactcat accaggcctg 6540
aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag ctttgttgta 6600
gggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacgggtctg cgttgctcggg 6660
aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac aaagccgccg 6720
tcccgtaagc tcagcgtaat gctctgccag tggtacaacc aattaaccaa ttctgattag 6780
aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt atcaatacca 6840
tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca gttccatagg 6900
atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat acaacctatt 6960
aatttccctc cgtcaaaaat aaggttatca agtgagaaat caccatgagt gacgactgaa 7020
tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac aggccagcca 7080
ttacgctcgt catcaaaatc actcgcacat accaaaccgt tattcattcg tgattgcgcc 7140
tgagcgagac gaaatacgcg atcgctgtta aaaggacaat tacaacagg aatcgaatgc 7200
aaccggcgca ggaacactgc cagcgcacat acaatatttt cacctgaatc aggatattct 7260
tctaatacct ggaatgctgt tttcccgggg atcgcagtg tgagtaacca tgcacatca 7320
ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgctcag ccagtttagt 7380
ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt cagaaacaac 7440
tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg cccgacatta 7500
tcgcgagccc atttataccc atataaatca gcacccatgt tggaatttaa tcgcggcctc 7560
gagcaagacg tttcccggtg aatatggctc ataacacccc ttgtattact gtttatgtaa 7620
gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta acatcagaga 7680
ttttgagaca caacgtggct tttccccccc cccattattt gaagcattta tcagggttat 7740
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 7800
cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 7860
acctataaaa ataggcgat cagcaggccc tttcgctc 7897

```

&lt;210&gt; 162

&lt;211&gt; 7984

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing SIV genes

&lt;400&gt; 162

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgct gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tgccccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720

```

catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccatga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtataca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctccg	cggccgggaa	cgggtgcattg	gaacgcggat	1020
tccccgtgcc	aagagtgacg	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctatagggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcc	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tattttttaca	1320
ggatgggggc	ccattttatta	tttacaattt	cacatataca	acaacgccgt	cccccggtcc	1380
cgcagttttt	attaaacata	gcgtgggac	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgcgcgc	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgacg	ccgccaccat	gaggccttgg	tccatgggga	aagaagcccc	1920
gcaatttccc	catggctcaa	gtgcatcagg	ggctgatgcc	aactgctccc	cgagaggacc	1980
cagctgtgga	tctgctaaag	aactacatgc	agttgggcaa	gcagcagaga	gaaaagcaga	2040
gagaaagcag	agagaagcct	tacaaggagg	tgacagagga	tttgctgcac	ctcaattctc	2100
tctttggagg	agaccagtag	tcactgcac	gatcgagggc	cagcccgtgg	aggtgcttct	2160
agacaccggc	gccgacgaca	gcacgtgac	cggcatcgag	ctgggcccgc	actacacccc	2220
gaagatcgtg	ggcggcatcg	gcggcttcat	caacaccgga	gagtacaaga	acgtggagat	2280
cgagggtgctg	ggcaagcgca	tcaagggcac	catcatgacc	ggcgacacgc	ccatcaacat	2340
cttcggccgc	aacctgctga	ccgccctggg	catgagcctg	aacttcccca	tcgccaaggt	2400
ggagcccgtg	aagggtggccc	tgaagcccgg	caaggacggc	cccaagctga	agcagtggcc	2460
cctgagcaag	gagaagatcg	tggccctgcg	cgagatctgc	gagaagatgg	agaaggacgg	2520
ccagctggag	gaggccccac	ccaccaaccc	ctacaacacg	cccaccttcg	ccatcaagaa	2580
gaaggacaag	aacaagtggc	gcacgtgat	cgacttccgc	gagctgaacc	gcgtgacca	2640
ggacttcacc	gaggtgcagc	tgggcatccc	gcaccccgc	ggcctggcca	agcgcaagcg	2700
catcaccgtg	ctggacatcg	gcgacgccta	cttcagcatc	cccctggacg	aggagtcccg	2760
ccagtacacc	gccttcaccc	tgccagcgt	gaacaacgcc	gagcccggca	agcgctacat	2820
ctacaaggtg	ctgccccagg	gctggaagg	cagcccgc	atcttccagt	acaccatgcg	2880
ccacgtgctg	gagcccttcc	gcaaggccaa	ccccgacgtg	accctgggtg	agtacatgca	2940
tgacatcctg	atcgccagcg	accgcaccga	cctggagcac	gaccgcgtgg	tgctgcagag	3000
caaggagctg	ctgaacagca	tcggcttcag	cacgcccag	gagaagttcc	agaaggaccc	3060
acccttccag	tggatgggct	acgagctgtg	gcccaccaag	tggaaagctg	agaagatcga	3120
gctgccccag	cgcgagacct	ggaccgtgaa	cgacatccag	aagctgggtg	gcgtgctgaa	3180
ctgggcccgc	cagatctacc	ccggcatcaa	gaccaagcac	ctgtgccgcc	tgatccgcgg	3240
caagatgacc	ctgaccgagg	aggtgcagtg	gaccgagatg	gccgaggccg	agtacgagga	3300
gaacaagatc	atcctgagcc	aggagcagga	gggctgctac	taccaggagg	gcaagccctt	3360
ggaggccacc	gtgatcaaga	gccaggacaa	ccagtggagc	tacaagatcc	accaggagga	3420
caagatcctg	aagggtgggca	agttcgccaa	gatcaagaac	acccacacca	acggcgtgcg	3480
cctgctggcc	cacgtgatcc	agaagatcgg	caaggaggcc	atcgtgatct	ggggccaggt	3540
gcccgaagttc	cacctgccc	tggagaagga	cgtgtgggag	cagtgggtgga	ccgactactg	3600
gcaggtgacc	tggataccgc	agtgggactt	catcagcacc	cctcccctgg	tgccctgggt	3660
gttcaacctg	gtgaaggacc	ccatcgaggg	cgaggagacc	tactacaccg	acggcagctg	3720
caacaagcag	agcaaggagg	gcaaggccgg	ctacatcacc	gaccgcggca	aggacaaggt	3780
gaagggtgctg	gagcagacca	ccaaccagca	ggccgagctc	gaggccttcc	tgatggccct	3840
gaccgacagc	ggccccaaag	ccaacatcat	cgtggacagc	cagtacgtga	tgggcatcat	3900
caccggctgc	cccaccgaga	gcgagagccg	cctgggtgaac	cagatcatcg	aggagatgat	3960
caagaagagc	gagatctacg	tggcctgggt	gcccgccac	aagggcacgc	gcggcaacca	4020



```

ggagatcgac cacctggtga gccagggcat ccgccagggtg ctgttcctgg agaagatcga 4080
gcccgeccag gaggagcacg acaagtacca cagcaacgtg aaggagctgg tgttcaagtt 4140
cggcctgccc cgcacgtggg cccgccagat cgtggacacc tgcgacaagt gccaccagaa 4200
gggcgaggcc atccacggcc aggccaacag cgacctgggc acctggcaga tggcctgcac 4260
ccacctggag ggcaagatca tcatcgtggc cgtgcacgtg gccagcggct tcatcgaggc 4320
cgaggtgatc ccgcaggaga ccggccgcca gaccgccctg ttctgctga agctggccgg 4380
ccgctggccc atcacccacc tgcacaccga caacggcgcc aacttcgcca gccaggaggt 4440
gaagatggtg gcctggtggg ccggcatcga gcacaccttc ggcgtgccct acaaccgcga 4500
gagccagggc gtggtggagg ccatgaacca ccacctgaag aaccagatcg accgcatccg 4560
cgagcaggcc aacagcgtgg agaccatcgt gctgatggcc gtgcactgca tgaacttcaa 4620
gcgcgcggc ggcatcggcg acatgacgcc cgccgagcgc ctgatcaaca tgatcaccac 4680
cgagcaggag atccagttcc agcagagcaa gaacagcaag ttcaagaact tccgcgtgta 4740
ctaccgcgag ggccgcgacc agctgtggaa gggccccggc gagctgctgt ggaagggcga 4800
gggcgcctg atcctgaagg tgggcaccga catcaagggtg gtgccccgcc gcaaggccaa 4860
gatcatcaag gactacggcg gcggcaagga ggtggacagc agcagccaca tggaggacac 4920
cggcgaggcc cgcgagggtg cctgagaatt cgtttaaaca cgcgtgtata cggatccaga 4980
tctgctgtgc cttctagttg ccagccatct gttgtttgcc cctccccgt gccttccttg 5040
accttggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcat 5100
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 5160
gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggg taccaggtg 5220
ctgaagaatt gacccggttc ctctggggc agaaagaagc aggcacatcc ccttctctgt 5280
gacacacct gtccacgccc ctggttctta gttccagccc cactcatagg aactcatag 5340
ctcaggaggg ctccgccttc aatcccaccc gctaaagtac ttggagcggg ctctccctcc 5400
ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta aagcaagata 5460
ggctattaag tgcagaggga gagaaaatgc ctccaacatg tgaggaagta atgagagaaa 5520
tcatagaatt ttaaggccat catggcctta atcttccgt tctcgtca ctgactcgt 5580
gcgctcggtc gttcggctgc ggcgagcggg atcagctcac tcaaaggcgg taatacgggt 5640
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 5700
caggaaccgt aaaaaggccg cgttgcctggc gtttttccat aggcctccgc cccctgacga 5760
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 5820
ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac 5880
cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 5940
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 6000
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtag 6060
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt 6120
aggcgggtgt acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt 6180
atthtggatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 6240
atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac 6300
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 6360
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 6420
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 6480
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 6540
tcgttcaccc atagtgcct gactcggggg gggggggcgc tgaggctctgc ctctgaaga 6600
agggtgtgct gactcatacc aggcctgaat cgcgccatca tccagccaga aagtgaggga 6660
gccacggttg atgagagctt tgtgtaggt ggaccagttg gtgattttga acttttgctt 6720
tgccacggaa cggctctgcgt tgcgggaag atgcgtgatc tgatccttca actcagcaaa 6780
agttcgattt attcaacaaa gccgcgctcc cgtcaagtca gcgtaatgct ctgccagtg 6840
tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat 6900
ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 6960
gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 7020
actcgtccaa catcaatata acctattaat ttccctcgt caaaaataag gttatcaagt 7080
gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 7140
ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcataacc 7200
aaaccgttat tcatcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa 7260
ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag cgcataaca 7320

```

```

atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc 7380
gcagtgggtga gtaacctatgc atcatcagga gtacggataa aatgcttgat ggtcggaaga 7440
ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 7500
ctacctttgc catgttttcag aaacaactct ggcgcatcgg gcttcccata caatcgatag 7560
attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata taaatcagca 7620
tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat atggctcata 7680
acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt 7740
ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc ccccccccc 7800
cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 7860
tagaaaaata aacaaatagg ggttcgcgcg acatttcccc gaaaagtgcc acctgacgtc 7920
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 7980
cgtc

```

&lt;210&gt; 163

&lt;211&gt; 9194

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 163

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcaggcgcg tcagcgggtg 120
ttggcgggtg tcggggcttg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtcctc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgcctt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860

```



tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcgagctgga 1920  
ccgctgggag aagatccgcc tgcgcccccg cggaagaag aagtacaagc tgaagcacat 1980  
cgtgtgggccc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040  
cgagggctgc cgccagatcc tgggcccagct gcagcccagc ctgcagaccg gcagcgagga 2100  
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160  
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220  
ggcccagcag gccgcccgcg acaccggcca cagcaaccag gtgagccaga actaccccat 2280  
cgtgcagaac atccagggcc agatggtgca ccaggccatc agcccccgca ccctgaacgc 2340  
ctgggtgaag gtggtggagg agaaggcctt cagccccgag gtgatcccca tggtcagcgc 2400  
cctgagcgag ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcc 2460  
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgcg agtgggaccg 2520  
cgtgcacccc gtgcacgcg gccccatcgc ccccgccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaacc 2640  
ccccatcccc gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700  
gcgcatgtac agccccacca gcatcctgga catccgccag ggccccaaagg agcccttccg 2760  
cgactacgtg gaccgcttct acaagacctt gcgcgcgag caggccagcc aggaggtgaa 2820  
gaactggatg accgagaccc tgctggtgca gaacgccaac cccgactgca agaccatcct 2880  
gaaggccctg ggccccgcg ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cggccccggc cacaaggccc gcgtgctggc cgaggccatg agccaggtga ccaacagcgc 3000  
caccatcatg atgcagcgc gcaacttccg caaccagcgc aagatcgtga agtgcttcaa 3060  
ctgcggcaag gagggccaca ccgcccgcaa ctgcccgcgc ccccgcaaga agggctgctg 3120  
gaagtgcggc aaggagggcc accagatgaa ggactgcacc gagcgacagg ctaatttttt 3180  
agggaagatc tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga 3240  
gccaacagcc ccaccagaag agagcttcag gtttggggaa gagacaacaa ctcccttca 3300  
gaagcaggag ccgatagaca aggaactgta tcctttagct tccctcagat cactcttgg 3360  
cagcgacccc tcgtcacat aaagataggg ggccagctga aggaggccct tctagacacc 3420  
ggcgcgcgac acaccgtgct ggaggagatg aacctgcccg gccgctggaa gccaagatg 3480  
atcggcggca tcggcggctt catcaagggt ggccagtacg accagatcct gatcgagatc 3540  
tgccggccaca aggccatcgg caccgtgctg gtgggcccc ccccgctgaa catcatcggc 3600  
cgcaacctgc tgaccagat cggctgcacc ctgaacttcc ccatcagccc catcgagacc 3660  
gtgcccgtga agctgaagcc cggcatggac ggccccaaagg tgaagcagtg gccctgacc 3720  
gaggagaaga tcaaggccct ggtggagatc tgcaccgaga tggagaagga gggcaagatc 3780  
agcaagatcg gccccgagaa cccctacaac acccccggtg tcgccatcaa gaagaaggac 3840  
agcaccaagt ggcgcaagct ggtggacttc cgcgagctga acaagcgcac ccaggacttc 3900  
tgggaggtgc agctgggcat cccccacccc gccggcctga agcagaagaa gagcgtgacc 3960  
gtgctggacg tgggcgacgc ctacttcagc gtgcccctgg acaaggactt ccgcaagtac 4020  
accgccttca ccatccccag catcaacaac gagaccccc gcacccgcta ccagtacaac 4080  
gtgctgcccc agggctggaa gggcagcccc gccatcttcc agtgcagcat gaccaagatc 4140  
ctggagccct tccgcaagca gaaccccgac atcgtgatct accagtacat ggacgacctg 4200  
tacgtgggca gcgacctgga gatcggccag caccgcacca agatcgagga gctgcgccag 4260  
cacctgctgc gctggggctt caccaccccc gacaagaagc accagaagga gcccccttc 4320  
ctgtggatgg gctacgagct gcacccccgac aagtggaccg tgcagcccat cgtgctgccc 4380  
gagaaggaca gctggaccgt gaacgacatc cagaagctgg tgggcaagct gaactgggcc 4440  
agccagatct acgccggcat caagggtgcgc cagctgtgca agctgctgcg cggcaccaag 4500  
gccctgaccg aggtggtgcc cctgaccgag gaggccgagc tggagctggc cgagaaccgc 4560  
gagatcctga aggagcccggt gcacggcgtg tactacgacc ccagcaagga cctgatcgcc 4620  
gagatccaga agcagggcca gggccagtgg acctaccaga tctaccagga gcccttcaag 4680  
aacctgaaga ccggcaagta cgcccgcatg aagggcgccc acaccaacga cgtgaagcag 4740  
ctgaccgagg ccgtgcagaa gatcgccacc gagagcatcg tgatctgggg caagaccccc 4800  
aagttcaagc tgcccatcca gaaggagacc tgggaggcct ggtggaccga gtactggcag 4860  
gccacctgga tccccgagtg ggagttcgtg aacaccccc ccctggtgaa gctgtggtac 4920  
cagctggaga aggagcccat catcggcgcc gagaccttct acgtggacgg cggcccaac 4980  
cgcgagacca agctgggcaa ggccggctac gtgaccgacc gcggccgcca gaagggtggtg 5040  
cccctgaccg acaccaccaa ccagaagacc gagctgcagg ccatccacct ggccctgcag 5100  
gacagcggcc tggaggtgaa catcgtgacc gacagccagt acgcccctggg catcatccag 5160

```

gcccagcccc acaagagcga gagegagctg gtgagccaga tcatcgagca gctgatcaag 5220
aaggagaagg tgtacctggc ctgggtgccc gccacaagg gcatcggcgg caacgagcag 5280
gtggacggcc tggtagcgc cggcatccgc aaggtgctgt tcctggacgg catcgacaag 5340
gcccaggagg agcacgagaa gtaccacagc aactggcgcg ccatggccag cgacttcaac 5400
ctgccccccg tggtagccaa ggagatcgtg gccagctgcg acaagtgcc gctgaagggc 5460
gaggccatgc acggccagggt ggactgcagc cccggcatct ggagctgga ctgcacccac 5520
ctggagggca aggtgatcct ggtggccgtg cacgtggcca gcggctacat cgaggccgag 5580
gtgatccccg ccgagaccgg ccaggagacc gcctacttcc tgctgaagct ggccggccgc 5640
tggcccgtag agaccgtgca caccgacaac ggcagcaact tcaccagcac caccgtgaag 5700
gccgcctgct ggtgggcccgg catcaagcag gagtccggca tcccctacaa ccccagagc 5760
cagggcgtga tcgagagcat gaacaaggag ctgaagaaga tcatcggcc ggtgcgagac 5820
caggccgagc acctgaagac cggcgtgcag atggccgtgt tcatccacaa cttcaagcgc 5880
aagggcggca tcggcggcta cagcgccggc gagcgcatcg tggacatcat cgccaccgac 5940
atccagacca aggagctgca gaagcagatc accaagatcc agaacttccg cgtgtactac 6000
cgcgacagcc gcgaccccgt gtggaagggc cccgccaaagc tgctgtgga gggcgagggc 6060
gccgtggtga tccaggacaa cagcgacatc aaggtggtgc cccgccgcaa ggccaagatc 6120
atccgcgact acggcaagca gatggccggc gacgactgcg tggccagccg ccaggacgag 6180
gactaggaat tctgctgtgc cttctagttg ccagccatct gttgtttgcc cctcccccg 6240
gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat 6300
tgcacgcgat tgtctgagta ggtgtcatte tattctgggg ggtgggggtg ggaggacag 6360
caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtg gctctatggg 6420
taccaggtg ctgaagaatt gaccgggttc ctctggggc agaaagaagc aggcacatcc 6480
ccttctctgt gacacaccct gtccacgccc ctggttctta gttccagccc cactcatagg 6540
acactcatag ctccaggagg ctccgccttc aatcccaccc gctaaagtac ttggagcgg 6600
ctctccctcc ctcatcagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta 6660
aagcaagata ggctattaag tgcagagggg gagaaaatgc ctccaacatg tgaggaaagta 6720
atgagagaaa tcatagaatt ttaaggccat catggcctta atcttccgct tcctcgctca 6780
ctgactcgct gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg 6840
taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 6900
agcaaaaggc caggaaccgt aaaaaggccg cggtgctggc gtttttccat aggtccgccc 6960
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 7020
tataaagata ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgaccc 7080
tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata 7140
gctcacgctg taggtatctc agttcggtgt aggtcggtcg ctccaagctg ggctgtgtgc 7200
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 7260
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 7320
cgaggtatgt aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 7380
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 7440
gtagctcttg atccggcaaa caaacaccg ctggtagcgg tggttttttt gtttgcaagc 7500
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacgggg 7560
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 7620
ggatcttcac ctagatcctt ttaaatataa aatgaagttt taaatcaatc taaagtatat 7680
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 7740
tctgtctatt tcgttcaccc atagttgcct gactcggggg gggggggcgc tgaggtctgc 7800
ctcgtgaaga aggtgttgct gactcatacc aggcctgaat cgcccatca tccagccaga 7860
aagtgaggga gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga 7920
acttttgctt tgccacggaa cggctcgcgt tgtcgggaag atgctgatc tgatccttca 7980
actcagcaaa agttcgatct attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct 8040
ctgccagtg tacaaccaat taaccaatc tgattagaaa aactcatcga gcatcaaatg 8100
aaactgcaat ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg 8160
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc 8220
tgcgattccg actcgtccaa catcaataca acctattaat tccccctcgt caaaaataag 8280
gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt 8340
atgcatttct tccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact 8400
cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgac 8460

```

```

gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 8520
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 8580
cccggggatc gcagtgggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 8640
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 8700
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata 8760
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 8820
taaatacagca tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat 8880
atggctcata acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 8940
tgatataatt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 9000
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 9060
tgaatgtatt tagaaaaata aacaaatagg ggtccgcgc acatttcccc gaaaagtgcc 9120
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 9180
gaggcccttt cgtc                                     9194

```

&lt;210&gt; 164

&lt;211&gt; 9194

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 164

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgccc accgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatth ccaagtcctc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatbg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttatbg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatatata acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800

```



cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctcttttc 1860  
tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcgagctgga 1920  
ccgctgggag aagatccgcc tgcgccccgg cggcaagaag aagtacaagc tgaagcacat 1980  
cgtgtgggccc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040  
cgagggctgc cgccagatcc tgggcccagct gcagcccagc ctgcagaccg gcagcgagga 2100  
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160  
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220  
ggcccagcag gccgcgcccg acaccggcca cagcaaccag gtgagccaga actaccccat 2280  
cgtgcagaac atccaggggc agatgggtgca ccaggccatc agcccccgca ccctgaacgc 2340  
ctgggtgaag gtgggtggagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgccc agtgggaccg 2520  
cgtgcacccc gtgcacgccc gccccatcgc cccgggccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaaccc 2640  
ccccatcccc gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700  
gcgcatgtac agccccacca gcatcctgga catccgcccag ggccccaaagg agcccttccg 2760  
cgactacgtg gaccgcttct acaagacctt gcgcgcccag caggccagcc aggaggtgaa 2820  
gaactggatg accgagaccc tgctgggtgca gaacgccaac cccgactgca agaccatcct 2880  
gaaggccctg ggccccgccc ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cgggccccggc cacaaggccc gcgtgctggc cgaggccatg agccagggtga ccaacagcgc 3000  
caccatcatg atgcagcgcg gcaacttccg caaccagcgc aagatcgtga agtgcttcaa 3060  
ctgcggcaag gaggggccaca ccgcccgcaa ctgcccgcgc ccccgcaaga agggctgctg 3120  
gaagtgcggc aaggaggggc accagatgaa ggactgcacc gagcgacagg ctaatttttt 3180  
agggaagatc tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga 3240  
gccaacagcc ccaccagaag agagcttcag gtttggggaa gagacaacaa ctccctctca 3300  
gaagcaggag ccgatagaca aggaactgta tccttttagct tccctcagat cactcttttg 3360  
cagcgacccc tcgtcacaat aaagataggg ggccagctga aggaggccct tctagacacc 3420  
ggcgccgacg acaccgtgct ggaggagatg aacctgcccg gccgctggaa gcccagatg 3480  
atcggcggca tcggcggtt catcaagggt ggccagtagc accagatcct gatcgagatc 3540  
tgcgccaca aggccatcgg caccgtgctg gtgggccccca ccccgctgaa catcatcggc 3600  
cgcaacctgc tgaccagat cggtgcacc ctgaacttcc ccatcagccc catcgagacc 3660  
gtgcccgtga agctgaagcc cggcatggac ggccccaaagg tgaagcagtg gccctgacc 3720  
gaggagaaga tcaaggccct ggtggagatc tgcaccgaga tggagaagga gggcaagatc 3780  
agcaagatcg gccccgagaa cccctacaac acccccgtgt tcgccatcaa gaagaaggac 3840  
agcaccaagt ggcgcaagct ggtggacttc cgcgagctga acaagcgcac ccaggacttc 3900  
tgggaggtgc agctgggcat cccccacccc gccggcctga agcagaagaa gagcgtgacc 3960  
gtgctggacg tgggcgacgc ctacttcagc gtgcccctgg acaaggactt ccgcaagtac 4020  
accgccttca ccatccccag catcaacaac gagacccccg gcatccgcta ccagtacaac 4080  
gtgctgcccc agggctggaa gggcagcccc gccatcttcc agtgagcat gaccaagatc 4140  
ctggagccct tccgcaagca gaaccccgac atcgtgatct accagtacat ggaccacctg 4200  
tacgtgggca gcgacctgga gatcggccag caccgcacca agatcgagga gctgcgccag 4260  
cacctgctgc gctggggctt caccaccccc gacaagaagc accagaagga gcccccttc 4320  
ctgtggatgg gctacgagct gcaccccgac aagtggaccg tgcagcccat cgtgctgccc 4380  
gagaaggaca gctggaccgt gaacgacatc cagaagctgg tgggcaagct gaactggggc 4440  
agccagatct acgccggcat caaggtgcgc cagctgtgca agctgctgcg cggcaccaag 4500  
gccctgaccg aggtgggtgcc cctgaccgag gaggccgagc tggagctggc cgagaaccgc 4560  
gagatcctga aggagcccg gcacggcgtg tactacgacc ccagcaagga cctgatcgcc 4620  
gagatccaga agcaggggcca gggccagtgg acctaccaga tctaccagga gcccttcaag 4680  
aacctgaaga ccggcaagta cgcccgcgtg aagggcgccc acaccaacga cgtgaagcag 4740  
ctgaccgagg ccgtgcagaa gatcgccacc gagagcatcg tgaatctggg caagaccccc 4800  
aagttcaagc tgcccatcca gaaggagacc tgggaggcct ggtggaccga gtactggcag 4860  
gccacctgga tccccgagtg ggagttcgtg aacaccccc ccctgggtgaa gctgtggtac 4920  
cagctggaga aggagcccat catcggcgcc gagaccttct acgtggacgg cgccgccaac 4980  
cgcgagacca agctgggcaa ggccggctac gtgaccgacc gcggccgcca gaagggtggtg 5040  
cccctgaccg acaccaccaa ccagaagacc gagctgcagg ccatccacct ggccctgcag 5100

gacagcggcc tggaggtgaa catcgtgacc gacagccagt acgccctggg catcatccag 5160  
gcccagcccg acaagagcga gagcgagctg gtgagccaga tcatcgagca gctgatcaag 5220  
aaggagaagg tgtacctggc ctgggtgccc gccacaaagg gcatcggcgg caacgagcag 5280  
gtggacggcc tggtagcgcc cggcatccgc aagggtgctgt tcctggacgg catcgacaag 5340  
gcccaggagg agcacgagaa gtaccacagc aactggcgcg ccatggccag cgacttcaac 5400  
ctgccccccg tggtagccaa ggagatcgtg gccagctgcg acaagtgcc a gctgaagggc 5460  
gaggccatgc acggccagggt ggactgcagc cccggcatct ggcagctggc atgcacccac 5520  
ctggagggca aggtgatect ggtggccgtg cacgtggcca gcggtacat cgaggccgag 5580  
gtgatccccg ccgagaccgg ccaggagacc gcctacttcc tgcgaagct ggccggccgc 5640  
tggcccgtag agaccgtgca caccgacaac ggcagcaact tcaccagcac caccgtgaag 5700  
gccgcctgct ggtgggcccgg catcaagcag gagttcggca tccctacaa ccccagagc 5760  
cagggcgtga tcgagagcat gaacaaggag ctgaagaaga tcatcgcca ggtgcgcgac 5820  
caggccgagc acctgaagac cggcgtgagc atggccgtgt tcatccaca cttcaagcgc 5880  
aagggcggca tcggcggcta cagcgccggc gagcgcatcg tggacatcat cgccaccgac 5940  
atccagacca aggagctgca gaagcagatc accaagatcc agaacttccg cgtgtactac 6000  
cgcgacagcc gcgaccccggt gtggaaggcc cccgccaagc tgctgtggaa gggcgagggc 6060  
gccgtggtga tccaggacaa cagcgacatc aagggtggtgc cccgcccga ggccaagatc 6120  
atccgcgact acggcaagca gatggccggc gacgactgcg tggccagccg ccaggacgag 6180  
gactaggaat tctgctgtgc cttctagttg ccagccatct gttgtttgcc cctccccgt 6240  
gccttccctg acctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat 6300  
tgcacgcat tgtctgagta ggtgtcatte tattctgggg ggtggggtgg ggcaggacag 6360  
caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggg 6420  
taccaggtg ctgaagaatt gaccgggttc ctctggggcc agaaagaagc aggcacatcc 6480  
ccttctctgt gacacacct gtccacgccc ctggttctta gttccagccc cactcatagg 6540  
acactcatag ctcaggagggt ctcgccttc aatcccaccc gctaaagtac ttggagcggc 6600  
ctctccctcc ctcacagcc caccaaacca aacctagcct ccaagagtgg gaagaaatta 6660  
aagcaagata ggctattaag tgcagagggg gagaaaatgc ctccaacatg tgaggaagta 6720  
atgagagaaa tcatagaatt ttaaggccat catggcctta atcttccgct tcctcgctca 6780  
ctgactcgct gcgctcggtc gttcggtgc ggcgagcggc atcagctcac tcaaaggcgg 6840  
taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 6900  
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggtccgcgc 6960  
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 7020  
tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgaccc 7080  
tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata 7140  
gctcacgctg taggtatctc agttcggtgt aggtcggttcg ctccaagctg ggctgtgtgc 7200  
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 7260  
accggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 7320  
cgaggtatgt aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 7380  
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 7440  
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 7500  
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg 7560  
ctgacgctca gtggaacgaa aactcacggt aagggtttt ggtcatgaga ttatcaaaaa 7620  
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 7680  
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 7740  
tctgtctatt tcgttcaccc atagttgcct gactcggggg gggggggcgc tgaggtctgc 7800  
ctcgtgaaga aggtgttgct gactcatacc aggcctgaat cgcccatca tccagccaga 7860  
aagtgagggg gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga 7920  
acttttgctt tgccacggaa cggctctgct tgtcgggaag atgcgtgac tgatccttca 7980  
actcagcaaa agttcgattt attcaacaaa gccgcgctcc cgtcaagtca gcgtaatgct 8040  
ctgccagtggt tacaaccaat taaccaatc tgattagaaa aactcatcga gcatcaaatg 8100  
aaactgcaat ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg 8160  
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc 8220  
tgcgattccg actcgtccaa catcaataca acctattaat tccctcgt caaaaataag 8280  
gttatcaagt gagaaatcac catgagtac gactgaatcc ggtgagaatg gcaaaagctt 8340  
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat caaatcact 8400



```

cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgac 8460
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag 8520
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt 8580
cccggggatc gcagtgggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat 8640
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc 8700
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata 8760
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata 8820
taaatacagca tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat 8880
atggctcata acacccttg tattactgtt tatgtaagca gacagtttta ttgttcatga 8940
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc 9000
ccccccccc cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 9060
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 9120
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 9180
gaggcccttt cgtc                                     9194

```

&lt;210&gt; 165

&lt;211&gt; 12411

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 165

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg gccattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740

```

gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca tgggcgcccc cgccagcgtg ctgagcggcg gcgagctgga 1920  
ccgctgggag aagatccgcc tgcgccccgg cggcaagaag aagtacaagc tgaagcacat 1980  
cgtgtgggccc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040  
cgagggctgc cgccagatcc tgggccagct gcagcccagc ctgcagaccg gcagcgagga 2100  
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160  
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220  
ggcccagcag gccgcgccc acaccggcca cagcaaccag gtgagccaga actaccccat 2280  
cgtgcagaac atccagggcc agatggtgca ccaggccatc agcccccgca ccctgaacgc 2340  
ctgggtgaag gtggtggagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgccc agtgggaccg 2520  
cgtgcacccc gtgcacgccc gcccacatcg ccccgccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gcaccctgca ggagcagatc ggctggatga ccaacaaccc 2640  
ccccatcccc gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700  
gcgcattgtac agccccacca gcatcctgga catccgccag ggccccaagg agcccttccg 2760  
cgactacgtg gaccgcttct acaagaccct gcgcgcccag caggccagcc aggaggtgaa 2820  
gaactggatg accgagaccc tgctggtgca gaacgccaac cccgactgca agaccatcct 2880  
gaaggccctg ggccccgccc ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cggccccggc cacaaggccc gcgtgctggc cgaggccatg agccaggtga ccaacagcgc 3000  
caccatcatg atgcagcgcg gcaacttccg caaccagcgc aagatcgtga agtgcttcaa 3060  
ctgcggcaag gagggccaca ccgccccgaa ctgcccgcgc ccccgcaaga agggctgctg 3120  
gaagtgcggc aaggagggcc accagatgaa ggactgcacc gagegacagg ctaatttttt 3180  
agggaagatc tggccttccc acaagggaag gccagggaat tttcttcaga gcagaccaga 3240  
gccaacagcc ccaccagaag agagcttcag gtttggggaa gagacaacaa ctccctctca 3300  
gaagcaggag ccgatagaca aggaactgta tcctttagct tccctcagat cactcttttg 3360  
cagcgacccc tcgtcacaat aaagataggg ggccagctga aggaggccct tctagacacc 3420  
ggcgccgacg acaccgtgct ggaggagatg aacctgcccg gccgctggaa gcccagatg 3480  
atcggcggca tcggcggtt catcaagggt ggccagtagc accagatcct gatcgagatc 3540  
tgccggccaca aggccatcgg caccgtgctg gtgggcccc ccccgctgaa catcatcggc 3600  
cgcaacctgc tgaccagat cggtgcacc ctgaacttcc ccatcagccc catcgagacc 3660  
gtgcccgtga agctgaagcc cggcatggac ggccccaagg tgaagcagtg gccctgacc 3720  
gaggagaaga tcaaggccct ggtggagatc tgcaccgaga tggagaagga gggcaagatc 3780  
agcaagatcg gccccgagaa cccctacaac acccccgtgt tcgccatcaa gaagaaggac 3840  
agcaccaagt ggcgcaagct ggtggacttc cgcgagctga acaagcgcac ccaggacttc 3900  
tgggaggtgc agctgggcat cccccacccc gccggcctga agcagaagaa gagcgtgacc 3960  
gtgctggacg tgggcgacgc ctacttcagc gtgcccctgg acaaggactt ccgcaagtac 4020  
accgccttca ccatccccag catcaacaac gagaccccc gcacccgcta ccagtacaac 4080  
gtgctgcccc agggctggaa gggcagcccc gccatcttcc agtgagcat gaccaagatc 4140  
ctggagccct tccgcaagca gaaccccgac atcgtgatct accagtacat ggacgacctg 4200  
tacgtgggca gcgacctgga gatcggccag caccgcacca agatcgagga gctgcgccag 4260  
cacctgctgc gctggggctt caccaccccc gacaagaagc accagaagga gcccccttc 4320  
ctgtggatgg gctacgagct gcaccccgac aagtggaccg tgcagcccat cgtgctgccc 4380  
gagaaggaca gctggaccgt gaacgacatc cagaagctgg tgggcaagct gaactgggccc 4440  
agccagatct acgcccgcac caagggtgcgc cagctgtgca agctgctgcg cggcaccaag 4500  
gccctgaccg aggtggtgcc cctgaccgag gaggccgagc tggagctggc cgagaaccgc 4560  
gagatcctga aggagcccgt gcacggcgctg tactacgacc ccagcaagga cctgatcgcc 4620  
gagatccaga agcagggcca gggccagtgg acctaccaga tctaccagga gcccttcaag 4680  
aacctgaaga ccggcaagta cgcccgcatg aaggcgccc acaccaacga cgtgaagcag 4740  
ctgaccgagg ccgtgcagaa gatcgccacc gagagcatcg tgatctgggg caagaccccc 4800  
aagttcaagc tgcccatcca gaaggagacc tgggaggcct ggtggaccga gtactggcag 4860  
gccacctgga tccccgagtg ggagttcgtg aacaccccc ccctggtgaa gctgtggtac 4920  
cagctggaga aggagcccat catcggcgcc gagaccttct acgtggacgg cgccgccaac 4980  
cgcgagacca agctgggcaa ggccggctac gtgaccgacc gcggccgcca gaagggtggtg 5040

ccccctgaccg acaccacca ccagaagacc gagctgcagg ccatccacct ggccctgcag 5100  
gacagcggcc tggaggtgaa catcgtgacc gacagccagt acgccctggg catcatccag 5160  
gcccagcccc acaagagcga gagcgagctg gtgagccaga tcatcgagca gctgatcaag 5220  
aaggagaagg tgtacctggc ctgggtgccc gccacaagg gcatcggcgg caacgagcag 5280  
gtggacggcc tggtagcgc cggcatccgc aaggtgctgt tcctggacgg catcgacaag 5340  
gcccaggagg agcacgagaa gtaccacagc aactggcgcg ccatggccag cgacttcaac 5400  
ctgccccccg tggtagccaa ggagatcgtg gccagctgcg acaagtgcc a gctgaagggc 5460  
gaggccatgc acggccaggt ggactgcagc cccggcatct ggcagctgga ctgcaccac 5520  
ctggagggca aggtgatcct ggtggccgtg cacgtggcca gcggctacat cgaggccgag 5580  
gtgatccccg ccgagaccgg ccaggagacc gcctacttcc tgctgaagct ggccggccgc 5640  
tggcccgtga agaccgtgca caccgacaac ggcagcaact tcaccagcac caccgtgaag 5700  
gccgcctgct ggtgggcccg catcaagcag gagttcggca tcccctacaa ccccagagc 5760  
cagggcgtga tcgagagcat gaacaaggag ctgaagaaga tcatcgcca ggtgcgcgac 5820  
caggccgagc acctgaagac cgccgtgcag atggccgtgt tcatccaca cttcaagcgc 5880  
aagggcggca tcggcggcta cagcgcggc gagcgcacgc tggacatcat cgccaccgac 5940  
atccagacca aggagctgca gaagcagatc accaagatcc agaacttccg cgtgtactac 6000  
cgcgacagcc gcgaccccgt gtggaagggc cccgccaaagc tgctgtggaa gggcgagggc 6060  
gccgtggtga tccaggacaa cagcgacatc aaggtggtgc cccgccgcaa ggccaagatc 6120  
atccgcgact acggcaagca gatggccggc gacgactgcg tggccagccg ccaggacgag 6180  
gactaggaat tcacgcgtcg agcatgcac tagggcggcc aattccgccc ctctccctcc 6240  
ccccccccta acgttactgg ccgaagccgc ttggaataag gccggtgtgc gtttgtctat 6300  
atgtgatttt ccaccatatt gccgtctttt ggcaatgtga gggcccgga acctggccct 6360  
gtcttcttga cgagcattcc taggggtctt tcccctctcg ccaaaggaat gcaaggtctg 6420  
ttgaatgtcg tgaaggaagc agttcctctg gaagcttctt gaagacaaac aacgtctgta 6480  
gcgacccttt gcaggcagcg gaacccccca cctggcgaca ggtgcctctg cgccaaaag 6540  
ccacgtgtat aagatacacc tgcaaaggcg gcacaacccc agtgccacgt tgtgagttgg 6600  
atagttgtgg aaagagtcaa atggctctcc tcaagcgtat tcaacaagg gctgaaggat 6660  
gcccagaagg taccctattg tatgggatct gatctggggc ctcggtgcac atgctttaca 6720  
tgtgtttagt cgaggttaaa aaaacgtcta ggccccccga accacgggga cgtggttttc 6780  
ctttgaaaaa cagcatgata agcttgccac aaccgggat cctctagaac accatgcgcg 6840  
tgaaggagaa gtaccagcac ctgtggcgct ggggctggcg ctggggcacc atgctgctgg 6900  
gcatgctgat gatctgcagc gccaccgaga agctgtgggt gaccgtgtac tacggcgtgc 6960  
ccgtgtggaa ggaggccacc accaccctgt tctgcgccag cgacgccaag gcctacgaca 7020  
ccgaggtgca caacgtgtgg gccaccacag cctgcgtgcc caccgacccc aacccccagg 7080  
aggtggtgct ggtgaacgtg accgagaact tcgacatgtg gaagaacgac atggtggagc 7140  
agatgcacga ggacatcatc agcctgtggg accagagcct gaagccctgc gtgaagctga 7200  
ccccctgtg cgtgagcctg aagtgcaccg acctgaagaa cgacaccaac accaacagca 7260  
gcagcggccg catgatcatg gagaaggcg agatcaagaa ctgcagcttc aacatcagca 7320  
ccagcatccg cggcaagggt cagaaggagt acgccttctt ctacaagctg gacatcatcc 7380  
ccatcgacaa cgacaccacc agctacagcc tgaccagctg caacaccagc gtgatcacc 7440  
aggcctgccc caaggtgagc ttcgagccca tcccctacca ctactgcgcc cccgccggct 7500  
tcgccatcct gaagtgaag gacaagaagt tcaacggcaa gggccctgc accaacgtga 7560  
gcaccgtgca gtgcaccac ggcatccgc ccgtggtgag caccagctg ctgctgaacg 7620  
gcagcctggc cgaggaggag gtggtgatcc gcagcgcaa ctctcgccgac aacgccaagg 7680  
tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc aacaacaaca 7740  
cccgcaagag catccacatc ggccccggcc gcgccttcta caccaccggc gagatcatcg 7800  
gcgacatccg ccaggccac tgcaacctga gccgcgcaa gtggaacgac accctgaaca 7860  
agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc aagcacagca 7920  
gcggcgggca ccccgagatc gtgaccacaa gcttcaactg cggcggcgag ttcttctact 7980  
gcaacagcac ccagctgttc aacagcacct ggttcaacag cacctggagc accgagggca 8040  
gcaacaacac cgagggcagc gacaccatca ccctgccctg ccgcatcaag cagatcatca 8100  
acatgtggca gaaggtgggc aaggccatgt acgcccccc catcagcggc cagatccgct 8160  
gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaacagc aacaacgaga 8220  
gcgagatctt ccgcctgggc ggcggcgaca tgcgcgacaa ctggcgagc gagctgtaca 8280  
agtacaaggt ggtgaagatc gagccccctg gcgtggcccc caccaaggcc aagcgccgcg 8340



tggtgcagcg	cgagaagcgc	gccgtgggca	tcggcgccct	gttcctgggc	ttcctgggcg	8400
ccgccggcag	caccatgggc	gccgccagca	tgacctgac	cgtgcaggcc	cgccagctgc	8460
tgagcggcat	cgtgcagcag	cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	8520
tgctgcagct	gaccgtgtgg	ggcatcaagc	agctgcaggc	ccgcctcctg	gccgtggagc	8580
gctacctgaa	ggaccagcag	ctgctgggca	tctggggctg	cagcggcaag	cttatctgca	8640
ccaccgccgt	gccctggaac	gccagctgga	gcaacaagag	cctggagcag	atctggaacc	8700
acaccacctg	gatggagtgg	gaccgcgaga	tcaacaacta	caccagcctg	atccacagcc	8760
tgatcgagga	gagccagaac	cagcaggaga	agaacgagca	ggagctgctg	gagctggaca	8820
agtgggccag	cctgtggaac	tggttcaaca	tcaccaactg	gctgtggtac	atcaagctgt	8880
tcacatgat	cgtgggcggc	ctggtgggcc	tgcgcctcgt	gttcgccgtg	ctgagcatcg	8940
tgaaccgcgt	gcgccagggc	tacagccccc	tgagcttcca	gacccacctg	cccatccccc	9000
gcggcccccga	ccgccccgag	ggcatcgagg	aggaggcgcg	cgagcgcgac	cgcgaccgca	9060
gcatccgcct	ggtgaacggc	agcctggccc	tgatctggga	cgacctgcgc	agcctgtgcc	9120
tgttcagcta	ccaccgcctg	cgcgacctgc	tgctgatcgt	gacccgcctc	gtggagctgc	9180
tggggccgccc	cggttgggag	gccctgaagt	actggtggaa	cctgctgcag	tactggagcc	9240
aggagctgaa	gaacagcgcc	gtgagcctgc	tgaacgccac	cgccatcgcc	gtggccgagg	9300
gcaccgaccg	cgtgatcgag	gtggtgcagg	gcgcctgccg	cgccatccgc	cacatccccc	9360
gccgcctccg	ccagggcctg	gagcgcctcc	tgctgtgagg	atccagatct	gctgtgcctt	9420
ctagttgcca	gccatctgtt	gtttgcccct	cccccgctgc	ttccttgacc	ctggaagggtg	9480
ccactcccac	tgctctttcc	taataaaaatg	aggaaattgc	atcgcatgtt	ctgagtaggt	9540
gtcattctat	tctgggggggt	ggggtggggc	aggacagcaa	gggggaggat	tgggaagaca	9600
atagcaggca	tgctggggat	gcggtgggct	ctatgggtac	ccagggtgctg	aagaattgac	9660
ccggttcctc	ctggggccaga	aagaagcagg	cacatccccct	tctctgtgac	acaccctgtc	9720
cacgccccctg	gttcttagtt	ccagcccccac	tcataaggaca	ctcatagctc	aggagggtctc	9780
cgccttcaat	cccacccgct	aaagtacttg	gagcggctctc	tccctccctc	atcagccccc	9840
caaaccaaac	ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	tattaagtgc	9900
agagggagag	aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	tagaatttta	9960
aggccatcat	ggccttaate	ttccgcttcc	tcgctcactg	actcgctgcg	ctcgggtcgtt	10020
cggctgcggc	gagcggatc	agctcactca	aaggcggtaa	tacggttatc	cacagaatca	10080
ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	10140
aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccccc	ctgacgagca	tcacaaaaat	10200
cgacgctcaa	gtcagagggtg	gcgaaacccg	acaggactat	aaagatacca	ggcgtttccc	10260
cctggaagct	ccctcgtgcg	ctctcctgtt	ccgaccctgc	cgcttacccg	atacctgtcc	10320
gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	10380
tcgggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aaccccccg	tcagccccgac	10440
cgctgcgcct	tatccggtaa	ctatcgtctt	gagtccaacc	cggtaagaca	cgacttatcg	10500
ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	cgggtgctaca	10560
gagttcttga	agtgggtggc	taactacggc	tacactagaa	gaacagtatt	tggtatctgc	10620
gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaaa	10680
accaccgctg	gtagcgggtg	tttttttgtt	tgcaagcagc	agattacgcg	cagaaaaaaa	10740
ggatctcaag	aagatccttt	gatcttttct	acggggctctg	acgctcagtg	gaacgaaaac	10800
tcacgttaag	ggatttttgt	catgagatta	tcaaaaaggga	tcttcacctc	gatectttta	10860
aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	agtaaacttg	gtctgacagt	10920
taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	gtctatcttc	ttcatccata	10980
gttgccctgac	tcgggggggg	ggggcgctga	ggtctgcctc	gtgaagaagg	tggtgctgac	11040
tcataccagg	cctgaatcgc	cccatcatcc	agccagaaag	tgagggagcc	acggttgatg	11100
agagctttgt	tgtaggtgga	ccagttgggtg	attttgaact	tttgctttgc	cacggaacgg	11160
tctgcgttgt	cgggaagatg	cgtgatctga	tccttcaact	cagcaaaagt	tcgattttat	11220
caacaaagcc	gccgtcccgt	caagtcagcg	taatgctctg	ccagtgttac	aaccaattaa	11280
ccaattctga	ttagaaaaac	tcatcgagca	tcaaatgaaa	ctgcaattta	ttcatatcag	11340
gattatcaat	accatatttt	tgaaaaagcc	gtttctgtaa	tgaaggagaa	aactcaccga	11400
ggcagttcca	taggatggca	agatcctggg	atcgggtctgc	gattccgact	cgtccaacat	11460
caatacaacc	tattaatttc	ccctcgtcaa	aaataagggt	atcaagtggg	aaatcaccat	11520
gagtgacgac	tgaatccggt	gagaatggca	aaagcttatg	catttctttc	cagacttggt	11580
caacaggcca	gccattacgc	tcgtcatcaa	aatcactcgc	atcaaccaa	ccgttattca	11640

```

ttcgtgattg cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa 11700
caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg 11760
aatcaggata ttcttctaata acctggaatg ctgttttccc ggggatcgca gtggtgagta 11820
accatgcac c atcaggagta cggataaaat gcttgatggt cggagagggc ataaattccg 11880
tcagccagtt tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat 11940
gtttcagaaa caactctggc gcacgagggt tcccatacaa tcgatagatt gtgcacacctg 12000
attgcccgcg attatcgcgga gcccatthtat acccatataa atcagcatcc atgttggaat 12060
ttaatcgcgg cctcgagcaa gacgtttccc gttgaatatg gctcataaca ccccttgat 12120
tactgtttat gtaagcagac agttttattg ttcattgatga tatattttta tcttggtgaa 12180
tgtaacatca gagattttga gacacaacgt ggctttcccc ccccccccat tattgaagca 12240
tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac 12300
aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta 12360
ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt c 12411

```

&lt;210&gt; 166

&lt;211&gt; 9170

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 166

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttccctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagtccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800

```



cgcgccacca	gacataatag	ctgacagact	aacagactgt	tccttttccat	gggtcttttc	1860
tgragtcacc	gtcgtcgaca	tgggcgcccc	cgccagcgtg	ctgagcggcg	gcgagctgga	1920
ccgcctgggag	aagatccgcc	tgcgcccccg	cggcaagaag	aagtacaagc	tgaagcacat	1980
cgtgtggggc	agccgcgagc	tggagcgctt	cgccgtgaac	cccgccctgc	tggagaccag	2040
cgagggctgc	cgccagatcc	tgggccagct	gcagcccagc	ctgcagaccg	gcagcgagga	2100
gctgcgcagc	ctgtacaaca	ccgtggccac	cctgtactgc	gtgcaccagc	gcatcgagat	2160
caaggacacc	aaggaggccc	tggacaagat	cgaggaggag	cagaacaaga	gcaagaagaa	2220
ggcccagcag	gccgccgccc	acaccggcca	cagcaaccag	gtgagccaga	actaccccat	2280
cgtgcagaac	atccagggcc	agatggtgca	ccaggccatc	agcccccgca	ccctgaacgc	2340
ctgggtgaag	gtgggtggagg	agaaggcctt	cagccccgag	gtgatcccca	tgttcagcgc	2400
cctgagcgag	ggcgccaccc	cccaggacct	gaacaccatg	ctgaacaccg	tgggcggcca	2460
ccaggccgcc	atgcagatgc	tgaaggagac	catcaacgag	gaggccgccc	agtgggaccg	2520
cgtgcacccc	gtgcacgccc	gccccatcgc	ccccggccag	atgcgcgagc	cccgcggcag	2580
cgacatcgcc	ggcaccacca	gcaccctgca	ggagcagatc	ggctggatga	ccaacaaccc	2640
ccccatcccc	gtgggcgaga	tctacaagcg	ctggatcatc	ctgggcctga	acaagatcgt	2700
gcgcattgtac	agccccacca	gcatcctgga	catccgccag	ggccccaagg	agcccttccg	2760
cgactacgtg	gaccgcttct	acaagaccct	gcgcgccgag	caggccagcc	aggaggtgaa	2820
gaactggatg	accgagaccc	tgctggtgca	gaacgccaac	cccgaactgc	agaccatcct	2880
gaaggccctg	ggccccgccc	ccaccctgga	ggagatgatg	accgcctgcc	agggcgtggg	2940
cggccccggc	cacaaggccc	gcgtgctggc	cgaggccatg	agccaggtga	ccaacagcgc	3000
caccatcatg	atgcagcgcg	gcaacttccg	caaccagcgc	aagatcgtga	agtgcctcaa	3060
ctgcggcaag	gagggccaca	ccgccccgaa	ctgcgcgcgc	ccccgcaaga	agggctgctg	3120
gaagtgcggc	aaggaggggc	accagatgaa	ggactgcacc	gagcgacagg	ctaataggga	3180
agatctggcc	ttcccacaag	ggaaggccag	ggaattttct	tcagagcaga	ccagagccaa	3240
cagccccacc	agaagagagc	ttcaggtttg	gggaagagac	aacaactccc	tctcagaagc	3300
aggagccgat	agacaaggaa	ctgtatcctt	tagcttccct	cagatcactc	tttggcagcg	3360
accctctgtc	acaataaaga	tagggggcca	gctgaaggag	gcccttctag	acaccggcgc	3420
cgacgacacc	gtgctggagg	agatgaacct	gccccggccc	tggaaagcca	agatgatcgg	3480
cggcatcggc	ggcttcatca	aggtggggcca	gtacgaccag	atcctgatcg	agatctgcgg	3540
ccacaaggcc	atcggcaccg	tgctggtggg	ccccaccccc	gtgaacatca	tggcccgcaa	3600
cctgctgacc	cagatcggct	gcaccctgaa	cttccccatc	agccccatcg	agaccgtgcc	3660
cgtgaagctg	aagccccgca	tggacggccc	caaggtgaag	cagtggcccc	tgaccgagga	3720
gaagatcaag	gccctggtgg	agatctgcac	cgagatggag	aaggaggggca	agatcagcaa	3780
gatcgccccc	gagaaccctt	acaacacccc	cgtgttcgcc	atcaagaaga	aggacagcac	3840
caagtggcgc	aagctggtgg	acttccgcga	gctgaacaag	cgcacccagg	acttctggga	3900
ggtgcagctg	ggcatcccc	accccgccgg	cctgaagcag	aagaagagcg	tgaccgtgct	3960
ggacgtgggc	gacgcctact	tcagcgtgcc	cctggacaag	gacttccgca	agtacaccgc	4020
cttcaccatc	cccagcatca	acaacgagac	ccccggcatc	cgctaccagt	acaacgtgct	4080
gccccagggc	tggaaaggga	gccccgccat	cttccagtgc	agcatgacca	agatcctgga	4140
gcccctccgc	aagcagaacc	ccgacatcgt	gatctaccag	tacatggacc	acctgtacgt	4200
gggcagcgac	ctggagatcg	gccagcaccg	caccaagatc	gaggagctgc	gccagcacct	4260
gctgcgctgg	ggcttcacca	cccccgacaa	gaagcaccag	aaggagcccc	ccttcctgtg	4320
gatgggctac	gagctgcacc	ccgacaagtg	gaccgtgcag	cccatcgtgc	tgcccagagaa	4380
ggacagctgg	accgtgaacg	acatccagaa	gctgggtggc	aagctgaact	gggcccagcca	4440
gatctacgcc	ggcatcaagg	tgcgccagct	gtgcaagctg	ctgcgcggca	ccaaggccct	4500
gaccgaggtg	gtgccccctga	ccgaggaggc	cgagctggag	ctggccgaga	accgcgagat	4560
cctgaaggag	cccgctgcacg	gcgtgtacta	cgaccccagc	aaggacctga	tgcgccgagat	4620
ccagaagcag	ggccagggcc	agtggacctt	ccagatctac	caggagccct	tcaagaacct	4680
gaagaccggc	aagtacgccc	gcatgaaggg	cgcccacacc	aacgacgtga	agcagctgac	4740
cgaggccgtg	cagaagatcg	ccaccgagag	catcgtgatc	tggggcaaga	cccccaagtt	4800
caagctgccc	atccagaagg	agacctggga	ggcctggtgg	accgagtact	ggcaggccac	4860
ctggatcccc	gagtgggagt	tctgtaacac	cccccccttg	gtgaagctgt	ggtaccagct	4920
ggagaaggag	cccatcatcg	gcgcgcgagac	cttctacgtg	gacggcgccc	ccaaccgcga	4980
gaccaagctg	ggcaaggccc	gctacgtgac	cgaccgcggc	cgccagaagg	tggtgccctt	5040
gaccgacacc	accaaccaga	agaccgagct	gcaggccatc	cacctggccc	tgcaggacag	5100

cggcctggag	gtgaacatcg	tgaccgacag	ccagtacgcc	ctgggcatca	tccaggcecca	5160
gcccgacaag	agcgagagcg	agctgggtgag	ccagatcatc	gagcagctga	tcaagaagga	5220
gaaggtgtac	ctggcctggg	tgcccgccca	caagggcatc	ggcggcaacg	agcaggtgga	5280
cggcctgggtg	agcgccggca	tccgcaaggt	gctgttcctg	gacggcatcg	acaaggccca	5340
ggaggagcac	gagaagtacc	acagcaactg	gcgcgccatg	gccagcgact	tcaacctgcc	5400
ccccgtgggtg	gccaaggaga	tcgtggccag	ctgcgacaag	tgccagctga	agggcgaggc	5460
catgcacggc	caggtggact	gcagcccccg	catctggcag	ctggcatgca	cccacctgga	5520
gggcaaggtg	atcctgggtg	ccgtgcacgt	ggccagcggc	tacatcgagg	ccgaggtgat	5580
ccccgccgag	accggccagg	agaccgccta	cttcctgctg	aagctggccg	gccgctggcc	5640
cgtgaagacc	gtgcacaccg	acaacggcag	caacttcacc	agcaccaccg	tgaaggccgc	5700
ctgctgggtg	gccggcatca	agcaggagtt	cggcatcccc	tacaaccccc	agagccaggg	5760
cgtgatcgag	agcatgaaca	aggagctgaa	gaagatcatc	ggccaggtgc	gcgaccaggc	5820
cgagcacctg	aagaccgccg	tgcagatggc	cgtgttcata	cacaacttca	agcgcaaggg	5880
cggcatcggc	ggctacagcg	ccggcgagcg	catcgtggac	atcatcgcca	ccgacatcca	5940
gaccaaggag	ctgcagaagc	agatcaccaa	gatccagaac	ttccgcgtgt	actaccgcga	6000
cagccgcgac	cccgtgtgga	agggccccgc	caagctgctg	tggaagggcg	agggcgccgt	6060
ggtgatccag	gacaacagcg	acatcaaggt	ggtgccccgc	cgcaaggcca	agatcatccg	6120
cgactacggc	aagcagatgg	ccggcgacga	ctgcgtggcc	agccgccagg	acgaggacta	6180
ggaattcttg	ctgtgccttc	tagttgccag	ccatctgttg	tttgccttc	ccccgtgcct	6240
tccttgaccc	tggaaggtgc	cactcccact	gtcctttcct	aataaaatga	ggaaattgca	6300
tcgcattgtc	tgagtaggtg	tcattctatt	ctgggggggtg	gggtggggca	gcacagcaag	6360
ggggaggatt	gggaagacaa	tagcaggcat	gctgggggatg	cggtgggctc	tatgggtacc	6420
caggtgctga	agaattgacc	cggttcctcc	tgggccagaa	agaagcaggc	acatccccct	6480
ctctgtgaca	cacctgtcc	acgccccctg	ttcttagtte	cagccccact	cataggacac	6540
tcatagctca	ggagggtctc	gccttcaatc	ccaccgccta	aagtacttgg	agcggctctc	6600
ccctccctca	tcagcccacc	aaaccaaacc	tagcctccaa	gagtgggaag	aaattaaagc	6660
aagataggct	attaagtgc	gaggagaga	aatgcctcc	aacatgtgag	gaagtaatga	6720
gagaaatcat	agaatttctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggctcgttcg	6780
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	6840
ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	6900
ggccgcgttg	ctggcgtttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	6960
acgctcaagt	cagaggtggc	gaaaccgcac	aggactataa	agataccagg	cgtttccccc	7020
tggaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	7080
ctttctccct	tcgggaagcg	tggcgctttc	tcaatgctca	cgctgtaggt	atctcagttc	7140
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	7200
ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	7260
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	7320
gttcttgaag	tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	7380
tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaac	7440
caccgctggt	agcgggtggt	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	7500
atctcaagaa	gatcctttga	tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	7560
acgttaaggg	attttgggtca	tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	7620
ttaaaaatga	agtttttaaat	caatctaaag	tatatatgag	taaacttggg	ctgacagtta	7680
ccaatgctta	atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	7740
tgctgactc	cggggggggg	gggcgctgag	gtctgcctcg	tgaagaagg	gttgctgact	7800
cataccaggc	ctgaatcgcc	ccatcatcca	gccagaaagt	gaggagacca	cgggtgatga	7860
gagctttggt	gtaggtggac	cagttggtga	ttttgaactt	ttgctttgcc	acggaacggt	7920
ctgcgttggt	gggaagatgc	gtgatctgat	ccttcaactc	agcaaaagtt	cgatttatct	7980
aacaaagccg	ccgtcccgtc	aagtcagcgt	aatgctctgc	cagtgttaca	accaattaac	8040
caattctgat	tagaaaaact	catcgagcat	caaatgaaac	tgcaatttat	tcatatcagg	8100
attatcaata	ccatattttt	gaaaaagccg	tttctgtaat	gaaggagaaa	actcaccgag	8160
gcagttccat	aggatggcaa	gacctgggtg	tcggtctgcg	attccgactc	gtccaacatc	8220
aatacaacct	attaatttcc	cctcgtcaaa	aataagggtg	tcaagtgaga	aatcaccatg	8280
agtgcgact	gaatccggtg	agaatggcaa	aagcttatgc	atttctttcc	agactgttcc	8340
aacaggccag	ccattacgct	cgtcatcaaa	atcactcgca	tcaaccaaac	cgttattcat	8400

```

tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac aattacaaac 8460
aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat ttccacctga 8520
atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgagc tggtagagtaa 8580
ccatgcatca tcaggagtag ggataaaatg cttgatgggc ggaagaggca taaattccgt 8640
cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac ctttgccatg 8700
tttcagaaac aactctggcg catcgggcct cccatacaat cgatagattg tcgcacctga 8760
ttgcccgcga ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt 8820
taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg ctcataacac cccttgatb 8880
actgtttatg taagcagaca gttttattgt tcatgatgat atatttttat cttgtgcaat 8940
gtaacatcag agattttgag acacaacgtg gctttccccc cccccccatt attgaagcat 9000
ttatcagggt tattgtctca tgagcgagata catatttgaa tgtatttaga aaaataaaca 9060
aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat 9120
tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 9170

```

&lt;210&gt; 167

&lt;211&gt; 9280

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing SIV genes

&lt;400&gt; 167

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
gggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgccg ggcattatgc ccagtacatg acctatggg actttccctac 660
ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggc tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtagc gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatgggggtc ccatttatta tttaaaaatt cacatatata acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcgagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggc ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800

```



cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgacg	ccgccaccat	gggcgtgcgc	aacagcgtgc	tgagcggcaa	1920
gaaggccgac	gagctggaga	agatccgcct	gcgcccgaac	ggcaagaaga	agtacatgct	1980
gaagcacgtg	gtgtgggccc	ccaacgagct	ggaccgcttc	ggcctggccg	agagcctgct	2040
ggagaacaag	gagggctgcc	agaagatcct	gagcgtgctg	gcgccgctgg	tgcccaccgg	2100
cagcgagaac	ctgaagagcc	tgtacaacac	cgtgtgcgtg	atctggtgca	tccacgccga	2160
ggagaagggtg	aagcacaccg	aggaggccaa	gcagatcgtg	cagcgccacc	tgggtggtgga	2220
gaccggcacc	accgagacca	tgcccagac	cagccgcccg	accgcgcccc	gcagcggacg	2280
cggcggcaac	tacccggtgc	agcagatcgg	cggcaactac	gtgcacctgc	cgctgagccc	2340
gcgcaccctg	aacgcctggg	tgaagctgat	cgaggagaag	aagttcggcg	ccgaggtggt	2400
gcccggcttc	caggccctga	gcgagggctg	cacgccctac	gacatcaacc	agatgctgaa	2460
ctgctgtggc	gaccaccagg	ccgccatgca	gatcatccgc	gacatcatca	acgaggaggc	2520
cgccgactgg	gacctgcagc	acccgcagcc	cgcgccgcag	cagggccagc	tgcgcgagcc	2580
cagcggcagc	gacatcgccc	gcaccaccag	cagcgtggac	gagcagatcc	agtggatgta	2640
ccgccagcag	aacccgatcc	cggtgggcaa	catctaccgc	cgctggatac	agctgggcct	2700
gcagaagtgc	gtgcgcatgt	acaacccgac	caacatcctg	gacgtgaagc	agggcccga	2760
ggagcccttc	cagagctacg	tggaccgctt	ctacaagagc	ctgcgcgccg	agcagaccga	2820
cgccgcccgtg	aagaactgga	tgaccagac	cctgctgata	cagaacgcc	acccggactg	2880
caagctggtg	ctgaagggcc	tgggcgtgaa	cccgaccctg	gaggagatgc	tgaccgcctg	2940
ccagggcgtg	ggcggcccgg	gccagaaggc	ccgcctgatg	gccgaggccc	tgaaggaggc	3000
cctggcgccc	gtgcccattc	cgttcgcggc	cgcccagcag	cgcgggcccgc	gcaagcccat	3060
caagtgtctg	aactgcggca	aggagggcca	cagcgcccgc	cagtgcgcgc	cgccgcgcgc	3120
ccagggctgc	tgggaagtgc	gcaagatgga	ccacgtgatg	gccaagtgcc	cggaccgcc	3180
ggcgggtagg	ccttggtcca	tggggaaaga	agccgcgcaa	tttcccgatg	gctcaagtgc	3240
atcaggggct	gatgccaaact	gctccgccag	aggaccagc	tgtggatctg	ctaaagaact	3300
acatgcagtt	gggcaagcag	cagagagaaa	agcagagaga	aagcagagag	aagccttaca	3360
aggaggtgac	agaggatttg	ctgcacctca	attctctctt	tggaggagac	cagtagtcac	3420
tgcacatgac	gagggccagc	ccgtggaggt	gcttctagac	accggcgccg	acgacagcat	3480
cgtgaccggc	atcgagctgg	gcccgcacta	caccccgaa	atcgtgggcg	gcacgcggcg	3540
cttcatcaac	accggagagt	acaagaacgt	ggagatcgag	gtgctgggca	agcgcatcaa	3600
gggcaccatc	atgaccggcg	acacgccc	caacatcttc	ggccgcaacc	tgctgaccgc	3660
cctgggcatg	agcctgaact	tccccatcgc	caaggtggag	cccgtgaagg	tggccctgaa	3720
gcccggcaag	gacggcccca	agctgaagca	gtggcccctg	agcaaggaga	agatcgtggc	3780
cctgcgcgag	atctgcgaga	agatggagaa	ggacggccag	ctggaggagg	ccccaccac	3840
caaccctac	aacacgcca	ccttcgccat	caagaagaag	gacaagaaca	agtggcgcat	3900
gctgatcgac	ttccgcgagc	tgaaccgcgt	gaccaggag	ttcaccgagg	tgacgtggg	3960
catcccgcac	cccgcgggcc	tggccaagcg	caagcgcate	accgtgctgg	acatcggcga	4020
cgcctacttc	agcatcccc	tggacgagga	gttccgccag	tacaccgcct	tcaccctgcc	4080
cagcgtgaac	aacgccgagc	ccggcaagcg	ctacatctac	aaggtgctgc	cccagggctg	4140
gaagggcagc	cccgccatct	tccagtacac	catgcgccac	gtgctggagc	ccttccgcaa	4200
ggccaacccc	gacgtgaccc	tgggtgcagta	catgcacgac	atcctgatcg	ccagcgaccg	4260
caccgacctg	gagcacgacc	gcgtgggtgt	gcagagcaag	gagctgctga	acagcatcgg	4320
cttcagcacg	cccgaggaga	agttccagaa	ggaccacccc	ttccagtgg	tgggctacga	4380
gctgtggccc	accaagtgga	agctgcagaa	gatcgagctg	ccccagcgcg	agacctggac	4440
cgtgaacgac	atccagaagc	tgggtggcgt	gctgaactgg	gccgcccaga	tctaccccgg	4500
catcaagacc	aagcacctgt	gccgcctgat	ccgcggcaag	atgaccctga	ccgaggaggt	4560
gcagtggacc	gagatggccg	aggccgagta	cgaggagaac	aagatcatcc	tgagccagga	4620
gcaggagggc	tgctactacc	aggagggcaa	gcccctggag	gccaccgtga	tcaagagcca	4680
ggacaaccag	tggagctaca	agatccacca	ggaggacaag	atcctgaagg	tgggcaagtt	4740
cgccaagatc	aagaacaccc	acaccaacgg	cgtgcgcctg	ctggcccacg	tgatccagaa	4800
gatcggcaag	gaggccatcg	tgatctgggg	ccaggtgccc	aagttccacc	tgcccgtgga	4860
gaaggacgtg	tgggagcagt	ggtggaccga	ctactggcag	gtgacctgga	tacccgagtg	4920
ggacttcate	agcacccttc	ccctggtgct	cctggtgttc	aacctggtga	aggaccccat	4980
cgagggcgag	gagacctact	acaccgacgg	cagctgcaac	aagcagagca	aggagggcaa	5040
ggccggctac	atcaccgacc	gcggcaagga	caaggtgaag	gtgctggagc	agaccaccaa	5100

ccagcaggcc gagctcgagg ccttcctgat ggccctgacc gacagcggcc ccaaggccaa 5160  
catcatcgtg gacagccagt acgtgatggg catcatcacc ggctgccccca ccgagagcga 5220  
gagccgcctg gtgaaccaga tcatcgagga gatgatcaag aagagcgaga tctacgtggc 5280  
ctgggtgccc gccacaagg gcatcggcgg caaccaggag atcgaccacc tggtagacca 5340  
gggcatccgc caggtgctgt tcctggagaa gatcgagccc gcccaggagg agcacgacaa 5400  
gtaccacagc aacgtgaagg agctgggtgt caagttcggc ctgccccgca tctgggcccg 5460  
ccagatcgtg gacacctgcg acaagtgcc aacagaaggc gaggccatcc acggccaggc 5520  
caacagcgac ctgggcacct ggcatatggc ctgcacccac ctggagggca agatcatcat 5580  
cgtggccgtg cacgtggcca gcggttcat cgaggccgag gtgatcccgc aggagaccgg 5640  
ccgccagacc gccctgttcc tgctgaagct ggccggccgc tggcccatca cccacctgca 5700  
caccgacaac ggccccaact tcgccagcca ggaggtgaag atggtggcct ggtgggcccg 5760  
catcgagcac accttcggcg tgccctacaa cccgcagagc cagggcgtgg tggaggccat 5820  
gaaccaccac ctgaagaacc agatcgaccg catccgcgag caggccaaca gcgtggagac 5880  
catcgtgctg atggccgtgc actgcattga cttcaagcgc cgcggcggca tcggcgacat 5940  
gacgcccgc gagcgctga tcaacatgat caccaccgag caggagatcc agttccagca 6000  
gagcaagaac agcaagttca agaacttccg cgtgtactac cgcgagggcc gcgaccagct 6060  
gtggaaggcc cccggcgagc tgctgtggaa gggcgaggcc gccgtgatcc tgaagggtgg 6120  
caccgacatc aagggtggtgc cccgccgcaa ggccaagatc atcaaggact acggcgccgg 6180  
caaggagggtg gacagcagca gccacatgga ggacaccggc gaggcccgcg aggtggcctg 6240  
agaattcgtt taaacacgcg tgtatacgga tccagatctg ctgtgccttc tagttgccag 6300  
ccatctgttg tttgcccctc ccccggtgct tccttgacct tggagggtgc cactcccact 6360  
gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt 6420  
ctgggggggtg ggggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat 6480  
gctggggatg cgggtgggctc tatgggtacc caggtgctga agaattgacc cggttcctcc 6540  
tggggccagaa agaagcaggc acatcccctt ctctgtgaca caccctgtcc acgcccctgg 6600  
ttcttagttc cagccccact cataggacac tcatagctca ggagggtccc gccttcaatc 6660  
ccaccgccta aagtacttgg agcgggtctc cctccctca tcagcccacc aaaccaaacc 6720  
tagcctccaa gagggtggaag aaattaaagc aagataggct attaatgca gaggagaga 6780  
aatgcctcc aacatgtgag gaagtaatga gagaaatcat agaattttaa ggccatcatg 6840  
gccttaattc tccgcttcc cgtcactga ctgcgtgcgc tcggctcgtc ggctgcggcg 6900  
agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc 6960  
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggcgcggtt 7020  
gctggcgttt tcccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag 7080  
tcagagggtg cgaaaccgca caggactata aagataccag gcgtttcccc ctggaagctc 7140  
cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc 7200  
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt 7260  
cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt 7320  
atccggtaac tatcgtcttg agtccaacce ggtaagacac gacttatcgc cactggcagc 7380  
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa 7440  
gtggtggcct aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa 7500  
gccagttacc ttcggaaaaa gaggttgtag ctcttgatcc ggcaaacaaa ccaccgctgg 7560  
tagcgggtgt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga 7620  
agatcctttg atcttttcta cggggtctga cgctcagtg aacgaaaact cacgttaagg 7680  
gattttgggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg 7740  
aagtttttaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt 7800  
aatcagttag gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact 7860  
cggggggggg gggcgctgag gtctgcctcg tgaagaagggt gttgctgact cataccaggc 7920  
ctgaatcgcc ccatcatcca gccagaaagt gaggagacca cggttgatga gagctttgtt 7980  
gtagggtggac cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc 8040  
gggaagatgc gtgatctgat ctttcaactc agcaaaagt cgatttatcc aacaaagccg 8100  
ccgtcccgtc aagtcagcgt aatgctctgc cagtggtaca accaatatc caattctgat 8160  
tagaaaaact catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata 8220  
ccatattttt gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat 8280  
aggatggcaa gatcctggta tcgggtctgcg attccgactc gtccaacatc aataaacct 8340  
attaatttcc cctcgtcaaa aataagggtt tcaagtgaga aatcaccatg agtgacgact 8400



```

gaatccggtg agaatggcaa aagcttatgc atttctttcc agacttggtc aacaggccag 8460
ccattacgct cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc 8520
gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa 8580
tgcaaccggc gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat 8640
tcttctaata cctggaatgc tgttttcccg gggatcgagc tggtagtaaa ccatgcatca 8700
tcaggagtac ggataaaatg cttgatggtc ggaagaggca taaattccgt cagccagttt 8760
agtctgacca tctcatctgt aacatcattg gcaacgctac ctttgccatg tttcagaaac 8820
aactctggcg catcgggctt cccatacaat cgatagattg tcgcacctga ttgcccagaca 8880
ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc 8940
ctcgagcaag acgtttcccg ttgaatatgg ctcataacac cccttgattt actgtttatg 9000
taagcagaca gttttattgt tcatgatgat atatttttat cttgtgcaat gtaacatcag 9060
agattttgag acacaacgtg gctttccccc ccccccat attgaagcat ttatcagggt 9120
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 9180
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 9240
ttaacctata aaaataggcg tatcacgagg ccctttcgtc 9280

```

&lt;210&gt; 168

&lt;211&gt; 9166

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 168

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagatttg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaat acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatatata acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccc cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```

gcagaagaag	atgcaggcag	ctgagttggt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgtaaacggt	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	tgggcgccc	cgccagcatc	ctgcgcggcg	gcaagctgga	1920
cgcctgggag	cgcataaagc	tgaagcccgg	cggcaagaag	cactacatga	tgaagcacct	1980
ggtgtggggc	agccgcgagc	tggagcgctt	cgccctggac	cccggcctgc	tggagaccag	2040
cgaggggctgc	aagcagatca	tgaagcagct	gcagcccgc	ctgcagaccg	gcaccaagga	2100
gctgatcagc	ctgcacaaca	ccgtggccac	cctgtactgc	gtgcacgaga	agatcgacgt	2160
gcgcgacacc	aaggaggccc	tggacaagat	caaggaggag	cagaacaaga	gccagcagaa	2220
gacccagcag	gccgaggccg	ccgacaaggg	caaggtgagc	cagaactacc	ccatcgtgca	2280
gaacctgcag	ggccagatgg	tgcaccagcc	catcagcgcc	cgcacctga	acgcctgggt	2340
gaaggtggtg	gaggagaagg	ccttcagccc	cgaggtgatc	cccatgttca	ccgcccctgag	2400
cgagggcgcc	acccccagg	acctgaacac	catgctgaac	accgtgggcg	gccaccaggc	2460
cgccatgcag	atgctgaagg	acaccatcaa	cgaggaggcc	gccgagtggg	accgcctgca	2520
ccccgtgcac	gccggccccg	tggcccccg	ccagatgcgc	gagccccgcg	gcagcgacat	2580
cgcgggcacc	accagcacc	tgcaggagca	gatcacctgg	atgaccaaca	accccccg	2640
gcccgtgggc	gacatctaca	agcgctggat	catcctgggc	ctgaacaaga	tcgtgcgc	2700
gtacagcccc	gtgagcatcc	tggacatcaa	gcagggcccc	aaggagccct	tcgcgcacta	2760
cgtggaccgc	ttcttcaaga	ccctgcgcgc	cgagcaggcc	accaggacg	tgaagaactg	2820
gatgaccgac	accctgctgg	tgcagaacgc	caaccccgac	tgcaagacca	tcctgcgcgc	2880
cctgggcccc	ggcgccagcc	tggaggagat	gatgaccgcc	tgccaggggcg	tgggcggccc	2940
cggccacaag	gcccgcgtgc	tggccgaggc	catgagcaag	gtgaacaaca	ccaacatcat	3000
gatgcagcgc	agcaactgca	agggccccaa	gcgaccatc	aagtgccttca	actgcggcaa	3060
ggagggccac	ctggccccgca	actgccgcgc	cccccgcaag	aagggtctgct	ggaagtgcgg	3120
caaggagggc	caccaggtga	aggactgcac	cgagcgacag	gctaattttt	tagggaagat	3180
ctggccttcc	cacaaggga	ggccaggga	tttcttcag	agcagaccag	agccaacagc	3240
cccaccagaa	gagagcttca	ggtttgggga	agagacaaca	actccctctc	agaagcagga	3300
gccgatagac	aaggaaactgt	atccttttagc	ttccctcaga	tcactctttg	gcagcgaccc	3360
ctcgtcacia	taaagatagg	gggcccagctg	aaggaggccc	ttctagacac	cggcgccgac	3420
gacaccgtgc	tggaggagat	gaacctgccc	ggcgctgga	agcccaagat	gatcggcggc	3480
atcggcggct	tcataaagg	gggcccagtac	gaccagatcc	tgatcgagat	ctgcggccac	3540
aaggccatcg	gcaccgtgct	ggtgggcccc	acccccgtga	acatcatcgg	ccgcaacctg	3600
ctgaccagaa	tcggctgcac	cctgaacttc	cccatcagcc	ccatcgagac	cgtgcccgtg	3660
aagctgaagc	ccggcatgga	cggccccaa	gtgaagcagt	ggcccctgac	cgaggagaag	3720
atcaaggccc	tgggtggagat	ctgcaccgag	atggagaagg	agggcaagat	cagcaagatc	3780
ggccccgaga	acccctacaa	cacccccgtg	ttcgccatca	agaagaagga	cagcaccaag	3840
tggcgcaagc	tgggtggactt	ccgcgagctg	aacaagcgca	cccaggactt	ctgggagggtg	3900
cagctgggca	tcacccaccc	cgccggcctg	aagcagaaga	agagcgtgac	cgtgctggac	3960
gtgggcgacg	cctacttcag	cgtgcccctg	gacaaggact	tcgcgaagta	caccgccttc	4020
accatcccca	gcatacaaca	cgagaccccc	ggcatccgct	accagtacaa	cgtgctgccc	4080
cagggtgga	agggcagccc	cgccatcttc	cagtgcagca	tgaccaagat	cctggagccc	4140
ttccgcaagc	agaacccccga	catcgtgata	taccagtaca	tggaccacct	gtacgtgggc	4200
agcgacctgg	agatcgggca	gcaccgcacc	aagatcgagg	agctgcgcca	gcacctgctg	4260
cgtggggct	tcaccacccc	cgacaagaag	caccagaagg	agccccctt	cctgtggatg	4320
ggctacgagc	tgcacccccga	caagtggacc	gtgcagccca	tcgtgctgcc	cgagaaggac	4380
agctggaccg	tgaacgacat	ccagaagctg	gtgggcaagc	tgaactgggc	cagccagatc	4440
tacgccggca	tcaagggtgcg	ccagctgtgc	aagctgctgc	gcggcaccac	ggccctgacc	4500
gaggtggtgc	ccctgaccga	ggaggccgag	ctggagctgg	ccgagaaccg	cgagatccctg	4560
aaggagcccc	tgcacggcgt	gtactacgac	cccagcaagg	acctgatcgc	cgagatccag	4620
aagcagggcc	agggccagtg	gacctaccag	atctaccagg	agcccttcaa	gaacctgaag	4680
accggcaagt	acgcccgcac	gaagggcgcc	cacaccaacg	acgtgaagca	gctgaccgag	4740
gccgtgcaga	agatcgccac	cgagagcate	gtgatctggg	gcaagacccc	caagttcaag	4800
ctgcccatacc	agaaggagac	ctgggaggcc	tgggtggaccg	agtactggca	ggccacctgg	4860
atccccgagt	gggagttcgt	gaacaccccc	cccctggtga	agctgtggta	ccagctggag	4920
aaggagccca	tcatacggcgc	cgagaccttc	tacgtggacg	gcgccgcca	ccgcgagacc	4980

```

aagctgggca aggccgggcta cgtgaccgac cgcggccgcc agaaggtggt gcccctgacc 5040
gacaccacca accagaagac cgagctgcag gccatccacc tggccctgca ggacagcggc 5100
ctggagggtga acatcgtgac cgacagccag tacgccctgg gcatcatcca ggcccagccc 5160
gacaagagcg agagcgagct ggtgagccag atcatcgagc agctgatcaa gaaggagaag 5220
gtgtacctgg cctgggtgcc cgcacacaag ggcacggcg gcaacgagca ggtggacggc 5280
ctggtgagcg ccggcatccg caaggtgctg ttcttgagc gcatcgacaa ggcccaggag 5340
gagcagcaga agtaccacag caactggcgc gccatggcca ggcacttcaa cctgcccccc 5400
gtggtggcca aggagatcgt ggccagctgc gacaagtgcc agctgaaggg cgaggccatg 5460
cacggccagg tggactgcag ccccggcac tggcagctgg catgcacca cctggagggg 5520
aaggtgatcc tgggtggcgt gcacgtggcc agcggctaca tcgaggccga ggtgatcccc 5580
gccgagaccg gccaggagac cgcctacttc ctgctgaagc tggccggccg ctggcccgctg 5640
aagaccgtgc acaccgacaa cggcagcaac ttcaccagca ccaccgtgaa ggccgcctgc 5700
tgggtgggccc gcatcaagca ggagtccggc atcccctaca acccccagag ccagggcgtg 5760
atcgagagca tgaacaagga gctgaagaag atcatcggcc aggtgcgcga ccaggccgag 5820
cacctgaaga ccgccgtgca gatggccgtg ttcattccaca acttcaagcg caaggcgggc 5880
atcgcgccgt acagcgccgg cgagcgcatc gtggacatca tcgccaccga catccagacc 5940
aaggagctgc agaagcagat caccaagatc cagaacttcc gcgtgtacta ccgcgacagc 6000
cgcgaccccg tgtggaaggg ccccgccaag ctgctgtgga agggcgaggg cgccgtggtg 6060
atccaggaca acagcgacat caaggtggtg ccccgccgca aggccaagat catccgcgac 6120
tacggcaagc agatggccgg cgacgactgc gtggccagcc gccaggacga ggactaggaa 6180
ttcttgctgt gccttctagt tgccagccat ctgttgtttg cccctcccc gtgccttcct 6240
tgaccctgga aggtgccact cccactgtcc ttctctaata aaatgaggaa attgcatcgc 6300
attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcagcac agcaaggggg 6360
aggattggga agacaatagc aggcattgtg gggatgcggg gggctctatg ggtaccag 6420
tgctgaagaa ttgaccgggt tctctctggg ccagaaagaa gcaggcacat ccccttctct 6480
gtgacacacc ctgtccacgc ccttggttct tagttccagc cccactcata ggacactcat 6540
agctcaggag ggctccgcct tcaatccac ccgctaaagt acttgagcg gtctctccct 6600
ccctcatcag cccaccaaac caaacctagc ctccaagagt ggggaagaaat taaagcaaga 6660
taggctatta agtgcagagg gagagaaaat gcctccaaca tgtgaggaag taatgagaga 6720
aatcatagaa tttcttccgc tctctcgtc actgactcgc tgcgctcggg cgttcggctg 6780
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 6900
gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcacia aaatcgacgc 6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 7020
agctccctcg tgcgctctcc tgttccgacc ctgcccgtta ccggatacct gtccgccttt 7080
ctcccttcgg gaagcgtggc gctttctcaa tgcctacgct gtaggtatct cagttcgggtg 7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 7200
gccttatccg gtaactatcg tcttgagtc aaccgggtaa gacacgactt atcgccactg 7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtg tacagagtgc 7320
ttgaagtggg ggccctaact cggctacact agaaggacag tatttggtat ctgcgctctg 7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 7440
gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 7500
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560
taagggatth tggatcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620
aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttggtctga cagttacca 7680
tgcttaataca gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc 7740
tgactccggg gggggggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata 7800
ccaggcctga atcgccccat catccagcca gaaagtgagg gagccacggg tgatgagagc 7860
tttggtgtag gtggaccagt tgggtgatttt gaacttttgc tttgccacgg aacggctctg 7920
gttgctcgga agatgcgtga tctgactctt caactcagca aaagttcgat ttattcaaca 7980
aagccgccgt cccgtcaagt cagcgtaatg ctctgcccag gttaacaacca attaaccaat 8040
tctgattaga aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta 8100
tcaataccat atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag 8160
ttccatagga tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata 8220
caacctatta atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg 8280

```

```

acgactgaat ccggtgagaa tggcaaaagc ttatgcattt ctttccagac ttgttcaaca 8340
ggccagccat tacgctcgtc atcaaaatca ctgcgcatcaa ccaaaccggt attcattcgt 8400
gattgcgcct gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga 8460
atcgaatgca accggcgagc gaacactgcc agcgcatcaa caatattttc acctgaatca 8520
ggatattctt ctaatacctg gaatgctgtt ttcccgggga tcgcagtggg gagtaaccat 8580
gcatcatcag gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc 8640
cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc 8700
agaaacaact ctggcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc 8760
ccgacattat cgcgagccca ttataccca tataaatcag catccatggt ggaatttaat 8820
cgcggcctcg agcaagacgt ttcccggtga atatggctca taacaccctt tgtattactg 8880
tttatgtaag cagacagttt tattgttcac gatgatatat ttttatcttg tgcaatgtaa 8940
catcagagat ttgagacac aacgtggctt tcccccccc cccattattg aagcatttat 9000
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 9060
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 9120
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 9166

```

&lt;210&gt; 169

&lt;211&gt; 9169

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 169

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aagacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaggt acgccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccct gagacgccat ccacgctggt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg ggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```



gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcaagtctga 1920  
cgccctgggag aagatccgcc tgcgccccgg cggcaagaag aagtaccgca tgaagcacct 1980  
gatctgggccc agccgcgagc tggaccgctt cgccctgaac ccagcctgc tggagaccgc 2040  
cgacggctgc cagcagatca tggagaagct gcagcccgcc ctgaagaccg gcaccgagga 2100  
gatcaagagc ctgttcaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgacgt 2160  
gaaggacacc aaggaggccc tggacaagat cgaggagatc aagaacaaga gcaagcagaa 2220  
gacccagcag gccgcccggc acaccggcaa cagcagcaac gtgagccaga actaccccat 2280  
cgtgcagaac gccagggccc agatgatcca ccagagcctg agcccccgca ccctgaacgc 2340  
ctgggtgaag gtgatcgagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc ccaggacct gaacatgatg ctgaacatcg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggacac catcaacgag gaggccggcg agtgggaccg 2520  
cctgcacccc gtgcacggcg gcccacccc ccccgccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gacccccca ggagcagatc ggctggatga ccggcaacc 2640  
ccccatcccc gtgggcgaca tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700  
gcgcattgtac agccccgtga gcatcctgga catcaagcag ggccccaaag agcccttccg 2760  
cgactacgtg gaccgcttct tcaagacctt ccgcgccgag caggccaccc aggacgtgaa 2820  
gaactggatg accgacaccc tgctgggtgca gaacgccaac ccgactgca aggccatcct 2880  
gcgcgccctg ggccccggcg ccaccctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cggccccggc cacaaggccc gcgtgctggc cgaggccatg agccaggtgc agcacaccaa 3000  
catcatgatg cagcgcggca acttcaaggg ccagaagcgc atcaagtgtc tcaactgcgg 3060  
caaggagggc cacctggccc gcaactgccg cgccctgcgc aagaagggt gctggaagt 3120  
cggcaaggag ggccaccaga tgaaggactg caccgagcga caggctaatt ttttagggaa 3180  
gatctggcct tcccacaagg gaaggccagg gaattttctt cagagcagac cagagccaac 3240  
agccccacca gaagagagct tcaggtttgg ggaagagaca acaactccct ctgagaagca 3300  
ggagccgata gacaaggaa tgtatccttt agcttccctc agatcactct ttggcagcga 3360  
ccctcgtca caataaagat agggggccag ctgaaggagg cccttctaga caccggcgcc 3420  
gacgacaccg tgctggagga gatgaacctg cccggccgct ggaagcccaa gatgatcggc 3480  
ggcatcggcg gcttcatcaa ggtgggcccag tacgaccaga tcctgatcga gatctgcggc 3540  
cacaaggcca tcggcacctg gctggtgggc cccacccccg tgaacatcat cggccgcaac 3600  
ctgctgaccc agatcggtg caccctgaac ttccccatca gcccacatga gaccgtgcc 3660  
gtgaagctga agcccgcat ggacggcccc aaggtgaagc agtggccct gaccgaggag 3720  
aagatcaagg ccctggtgga gatctgcacc gagatggaga aggagggcaa gatcagcaag 3780  
atcgcccccg agaacccta caacaccccc gtgttcgcca tcaagaagaa ggacagcacc 3840  
aagtggcgca agctggtgga ctccgcgag ctgaacaagc gcaccagga cttctgggag 3900  
gtgcagctgg gcatcccca ccccgccggc ctgaagcaga agaagagcgt gaccgtgctg 3960  
gacgtgggcg acgcctactt cagcgtgccc ctggacaagg acttccgcaa gtacaccgcc 4020  
ttcaccatcc ccagcatcaa caacgagacc cccggcatcc gctaccagta caacgtgctg 4080  
ccccagggtt ggaagggcag ccccgccatc ttccagtga gcatgaccaa gatcctggag 4140  
cccttccgca agcagaaccc cgacatcgtg atctaccagt acatggacca cctgtacgtg 4200  
ggcagcgacc tggagatcgg ccagcaccgc accaagatcg aggagctgcg ccagcacctg 4260  
ctgcgctggg gcttcaccac ccccgacaag aagcaccaga aggagcccc cttcctgtgg 4320  
atgggctacg agctgcacc cgacaagtgg accgtgcagc ccacgtgct gcccgagaag 4380  
gacagctgga ccgtgaacga catccagaag ctggtgggca agctgaactg ggccagccag 4440  
atctacgccc gcatcaaggc gcgccagctg tgcaagctgc tgcgcggcac caaggccctg 4500  
accgaggtgg tgccctgac cgaggaggcc gagctggagc tggccgagaa ccgcgagatc 4560  
ctgaaggagc ccgtgcacgg cgtgtactac gacccagca aggacctgat cgccgagatc 4620  
cagaagcagg gccagggcca gtggacctac cagatctacc aggagccctt caagaacctg 4680  
aagaccggca agtacgccc catgaagggc gccacacca acgacgtgaa gcagctgacc 4740  
gaggccgtgc agaagatcg caccgagagc atcgtgatct ggggcaagac cccaagtctc 4800  
aagctgcca tccagaagga gacctgggag gcctgggtgga ccgagtactg gcaggccacc 4860  
tggatccccg agtgggagtt cgtgaacacc cccccctgg tgaagctgtg gtaccagctg 4920  
gagaaggagc ccatcatcgg cgccgagacc ttctacgtgg acggcgccgc caaccgcgag 4980



accaagctgg gcaaggccgg ctacgtgacc gaccgcggcc gccagaaggt ggtgcccctg 5040  
accgacacca ccaaccagaa gaccgagctg caggccatcc acctggccct gcaggacagc 5100  
ggcctggagg tgaacatcgt gaccgacagc cagtacgccc tgggcatcat ccaggcccag 5160  
cccgacaaga gcgagagcga gctggtgagc cagatcatcg agcagctgat caagaaggag 5220  
aaggtgtacc tggcctgggt gcccgcccac aagggcacg gcggcaacga gcaggtggac 5280  
ggcctggtga gcgccggcat ccgcaagggt ctgttcctgg acggcatcga caaggcccag 5340  
gaggagcacg agaagtacca cagcaactgg cgcgccatgg ccagcgactt caacctgccc 5400  
cccgtggtgg ccaaggagat cgtggccagc tgcgacaagt gccagctgaa gggcgaggcc 5460  
atgcacggcc aggtggactg cagccccggc atctggcagc tggcatgcac ccacctggag 5520  
ggcaagggtga tcctggtggc cgtgcacgtg gccagcggct acatcgaggc cgaggtgate 5580  
cccgccgaga ccggccagga gaccgcctac ttctgtctga agctggccgg ccgctggccc 5640  
gtgaagaccg tgcacaccga caacggcagc aacttcacca gcaccaccgt gaaggccgcc 5700  
tgctggtggg ccggcatcaa gcaggagttc ggcattccct acaacccccca gagccagggc 5760  
gtgatcgaga gcatgaacaa ggagctgaag aagatcatcg gccaggtgcy cgaccaggcc 5820  
gagcacctga agaccgccgt gcagatggcc gtgttcattc acaacttcaa gcgcaagggc 5880  
ggcatcggcg gctacagcgc cggcgagcgc atcgtggaca tcatcgccac cgacatccag 5940  
accaaggagc tgcagaagca gatcaccaag atccagaact tccgcgtgta ctaccgcgac 6000  
agccgcgacc ccgtgtggaa gggccccgcc aagctgctgt ggaagggcga gggcgccgtg 6060  
gtgatccagg acaacagcga catcaagggt gtgccccgcc gcaaggccaa gatcatccgc 6120  
gactacggca agcagatggc cggcgacgac tgcgtggcca gccgccagga cgaggactag 6180  
gaattcttgc tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt 6240  
ccttgaccct ggaagggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat 6300  
cgcattgtct gagtaggtgt cattctattc tgggggggtg ggtggggcag cacagcaagg 6360  
gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atgggtaccc 6420  
aggtgctgaa gaattgacct ggttcctcct gggccagaaa gaagcaggca catccccttc 6480  
tctgtgacac accctgtcca cggccctggt tcttagttcc agccccactc ataggacact 6540  
catagctcag gagggctccg ctttcaatcc caccgcctaa agtacttgga gcggtctctc 6600  
cctccctcat cagcccacca aaccaaactc agcctccaag agtggaaga aattaaagca 6660  
agataggcta ttaagtgcag agggagagaa aatgcctcca acatgtgagg aagtaatgag 6720  
agaaatcata gaatttcttc cgtctcctcg ctactgact cgctgcgctc ggtcgttcgg 6780  
ctgcggcgag cgttatcagc tcaactcaaag gcggtaatat ggttatccac agaatcaggg 6840  
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 6900  
gccgcgttgc tggcggtttt ccataggctc cggccccctg acgagcatca caaaaatcga 6960  
cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 7020  
ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 7080  
tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg 7140  
gtgtaggteg ttcgctccaa gctgggctgt gtgcacgaac ccccggttca gcccgaccgc 7200  
tgcgccttat ccggtacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 7260  
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 7320  
ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 7380  
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc 7440  
accgctggta gcggtgggtt ttttgggttc aagcagcaga ttacgcgag aaaaaaagga 7500  
tctcaagaag atcctttgat cttttctacg ggggtctgacg ctcagtggaa cgaaaactca 7560  
cgtaaggga ttttggctcat gagattatca aaaaggatct tcacctagat ctttttaaatt 7620  
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 7680  
caatgcttaa tcagtaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 7740  
gcctgactcc gggggggggg ggcgctgagg tctgcctcgt gaagaagggt ttgctgactc 7800  
ataccaggcc tgaatcgccc catcatccag ccagaaagtg agggagccac ggttgatgag 7860  
agctttgttg taggtggacc agttggtgat tttgaacttt tgctttgcca cggaacggtc 7920  
tgcggtgtcg ggaagatgcg tgatctgac cttcaactca gcaaaagttc gatttattca 7980  
acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtggttaca ccaattaacc 8040  
aattctgatt agaaaaactc atcgagcabc aaatgaaact gcaatttatt catatcagga 8100  
ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg 8160  
cagttccata ggatggcaag atcctggtat cggctctgca ttccgactcg tccaacatca 8220  
atacaaccta ttaatttccc ctgcgtcaaaa ataagggtat caagtgagaa atcaccatga 8280

```

gtgacgactg aatccggtga gaatggcaaa agcttatgca tttctttcca gacttggttca 8340
acaggccagc cattacgctc gtcatacaaaa tcaactcgcat caaccaaacc gttattcatt 8400
cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca attacaaaca 8460
ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa 8520
tcaggatatt cttctaatac ctggaatgct gttttcccg ggatcgcagt ggtgagtaac 8580
catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgtc 8640
agccagttta gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt 8700
ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt cgcacctgat 8760
tgcccgacat tatcgcgagc ccattttatac ccataataat cagcatccat gttggaattt 8820
aatcgcggcc tcgagcaaga cgtttcccg tgaatatggc tcataacacc ccttgattta 8880
ctgtttatgt aagcagacag ttttattgtt catgatgata tatttttatc ttgtgcaatg 8940
taacatcaga gattttgaga cacaacgtgg ctttcccccc cccccatta ttgaagcatt 9000
tatcagggtt attgtctcat ggcgggatac atatttgaat gtatttagaa aaataaaca 9060
ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt 9120
atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc 9169

```

&lt;210&gt; 170

&lt;211&gt; 9792

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 170

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgccctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcc tggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgcg tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgct agatcgccct gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatgt gaacgcggat 1020
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccattttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgtcgg cagctccttg ctcttaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680

```

gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcgtg ctgagcggcg gcgagctgga 1920  
ccgctgggag aagatccgcc tgcgccccg cggaagaag aagtacaagc tgaagcacat 1980  
cgtgtgggcc agccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagaccag 2040  
cgagggctgc cgccagatcc tgggccagct gcagcccagc ctgcagaccg gcagcgagga 2100  
gctgcgcagc ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgagat 2160  
caaggacacc aaggaggccc tggacaagat cgaggaggag cagaacaaga gcaagaagaa 2220  
ggcccagcag gccgcccgcg acaccggcca cagcaaccag gtgagccaga actaccccat 2280  
cgtgcagaac atccagggcc agatggtgca ccaggccatc agcccccgca ccctgaacgc 2340  
ctgggtgaag gtggtggagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc ccaggacct gaacaccatg ctgaacaccg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggagac catcaacgag gaggccgccg agtgggaccg 2520  
cgtgcacccc gtgcacgccc gccccatcgc cccgggccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gcacctgca ggagcagatc ggctggatga ccaacaacc 2640  
ccccatcccc gtgggcgaga tctacaagcg ctggatcctc ctgggcctga acaagatcgt 2700  
gcgcatgtac agccccacca gcattctgga catccgccag ggccccagg agcccttccg 2760  
cgactacgtg gaccgcttct acaagacct gcgcgccgag caggccagcc aggaggtgaa 2820  
gaactggatg accgagaccc tgctggtgca gaacgccaac cccgactgca agaccatcct 2880  
gaaggccctg ggcgccgccc ccacctgga ggagatgatg accgcctgcc agggcgtggg 2940  
cgcccccgcc cacaaggccc gcgtgctggc cgaggccatg agccaggtga ccaacagcgc 3000  
caccatcatg atgcagcgcg gcaacttccg caaccagcgc aagatcgtga agtgcttcaa 3060  
ctgcggcaag gagggccaca ccgcccgcga ctgccgcgcc ccccgcaaga agggctgctg 3120  
gaagtgcggc aaggagggcc accagatgaa ggactgcacc gagcgacagg ctaataggga 3180  
agatctggcc tccccacaag ggaaggccag ggaattttct tcagagcaga ccagagccaa 3240  
cagccccacc agaagagagc ttcaggtttg gggaagagac acaactccc tctcagaagc 3300  
aggagccgat agacaaggaa ctgtatcctt tagcttccct cagatcactc tttggcagcg 3360  
acctctcgtc acaataaaga tagggggcca gctgaaggag gcccttctag acaccggcgc 3420  
cgacgacacc gtgctggagg agatgaacct gccgggccgc tggaagccca agatgatcgg 3480  
cggcatcggc ggcttcatca aggtgggcca gtacgaccag atcctgatcg agatctgcgg 3540  
ccacaaggcc atcggcaccg tgctggtggg cccaccccc gtgaacatca tcggccgcaa 3600  
cctgctgacc cagatcggct gcacctgaa cttccccatc agccccatcg agaccgtgcc 3660  
cgtgaagctg aagcccggca tggacggccc caaggtgaag cagtggcccc tgaccgagga 3720  
gaagatcaag gccctggttg agatctgcac cgagatggag aaggagggca agatcagcaa 3780  
gatcggcccc gagaacctt acaacacccc cgtgttcgcc atcaagaaga aggacagcac 3840  
caagtggcgc aagctggttg acttccgcga gctgaacaag cgcaccagc acttctggga 3900  
ggtgcagctg ggcaccccc accccgcccg cctgaagcag aagaagagcg tgaccgtgct 3960  
ggacgtgggc gacgcctact tcagcgtgcc cctggacaag gacttccgca agtacaccgc 4020  
cttcaccatc cccagcatca acaacgagac cccggcatc cgctaccagt acaacgtgct 4080  
gccccagggc tggaggggca gcccgcctat cttccagtgc agcatgacca agatcctgga 4140  
gcccttccgc aagcagaacc ccgacatcgt gatctaccag tacatggacc acctgtacgt 4200  
gggcagcgac ctggagatcg gccagcaccg caccaagatc gaggagctgc gccagcacct 4260  
gctgcgctgg ggcttcacca ccccgacaa gaagcaccag aaggagcccc ccttctgtg 4320  
gatgggctac gagctgcacc ccgacaagt gaccgtgcag cccatcgtgc tggccgagaa 4380  
ggacagctgg accgtgaacg acatccagaa gctggtgggc aagctgaact gggccagcca 4440  
gatctacgcc ggcataaagg tgcgccagct gtgcaagctg ctgcgcggca ccaaggccct 4500  
gaccgaggtg gtgcccctga ccgaggaggc cgagctggag ctggccgaga accgcgagat 4560  
cctgaaggag cccgtgcacg gcgtgtacta cgaccccagc aaggacctga tcgccgagat 4620  
ccagaagcag ggccagggcc agtggacctt ccagatctac caggagccct tcaagaacct 4680  
gaagaccggc aagtacgcc gcataaggg cgccacacc aacgacgtga agcagctgac 4740  
cgaggccgtg cagaagatcg ccaccgagag catcgtgatc tggggcaaga ccccaagtt 4800  
caagctgcc atccagaagg agacctggga ggctggtgg accgagtact ggcaggccac 4860  
ctggatcccc gagtgggagt tcgtgaacac cccccctg gtgaagctgt ggtaccagct 4920  
ggagaaggag cccatcatcg gcgcgagac cttctacgtg gacggcgccg ccaaccgcga 4980



gaccaagctg ggcaaggccg gctacgtgac cgaccgcggc cgccagaagg tgggtgcccct 5040  
gaccgacacc accaaccaga agaccgagct gcaggccatc cacctggccc tgcaggacag 5100  
cggcctggag gtgaacatcg tgaccgacag ccagtaacgc ctggggcatca tccaggccca 5160  
gcccgacaag agcgagagcg agctgggtgag ccagatcatc gagcagctga tcaagaagga 5220  
gaagggtgtac ctggcctggg tgcccgccca caagggcacg ggcggaacg agcagggtga 5280  
cggcctgggt agcgccggca tccgcaaggt gctgttcctg gacggcatcg acaaggccca 5340  
ggaggagcac gagaagtacc acagcaactg gcgcgccatg gccagcgact tcaacctgcc 5400  
ccccgtgggt gccaaaggaga tcgtggccag ctgcgacaag tgccagctga agggcgaggc 5460  
catgcacggc cagggtggact gcagccccgg catctggcag ctggcatgca cccacctgga 5520  
gggcaagggt atcctgggtg ccgtgcacgt ggccagcggc tacatcgagg ccgagggtgat 5580  
ccccgccgag accggccagg agaccgccta ctctctgctg aagctggccg gccgctggcc 5640  
cgtgaagacc gtgcacaccg acaacggcag caacttcacc agcaccaccg tgaaggccgc 5700  
ctgctgggtg gccggcatca agcaggagtt cggcatcccc tacaaccccc agagccaggg 5760  
cgtgatcgag agcatgaaca aggagctgaa gaagatcatc ggccagggtg gcgaccaggc 5820  
cgagcacctg aagaccgccg tgcagatggc cgtgttcac cacaacttca agcgcaaggg 5880  
cggcatcggc ggctacagcg ccggcgagcg catcgtggac atcatcgcca ccgacatcca 5940  
gaccaaggag ctgcagaagc agatcaccaa gatccagaac ttccgcgtgt actaccgcga 6000  
cagccgcgac cccgtgtgga agggccccgc caagctgctg tggaaggggc agggcgccgt 6060  
ggtgatccag gacaacagcg acatcaaggt ggtgccccgc cgcaaggcca agatcatccg 6120  
cgactacggc aagcagatgg ccggcgacga ctgctggcc agccgccagg acgaggacat 6180  
gggcggaag tggagcaaga gcagcgtgat cggctggccc gccgtgcgcg agcgcatgcg 6240  
ccgcgccgag cccgccgcg acggcgtggg cggcgtgagc cgcgacctgg agaagcacgg 6300  
cgccatcacc agcagcaaca ccgccgcaa caacgccgc tgccctggc tggaggccca 6360  
ggaggaggag gaggtgggt tccccgtgac ccccagggtg cccctgcgc ccctgacctt 6420  
caaggccgcc gtggacctga gccacttcct gaaggagaag ggcggcctgg agggcctgat 6480  
ccacagccag cgcgccagg acatcctgga cctgtggatc taccacacc agggctactt 6540  
ccccgactgg cagaactaca ccccgggccc cggcgtgcgc taccctctga ccttcggctg 6600  
gtgctacaag ctgggtgccc tggagcccg caaggtggag gaggccaaca agggcgagaa 6660  
caccagcctg ctgcacccc tgagcctgca cggcatggac gaccccgagc gcgagggtgt 6720  
ggagtggcgc ttccgacagc gcctggcctt ccaccacgtg gcccgcgagc tgcacccga 6780  
gtacttcaag aactgctgag gatcaattct tgctgtgcct tctagtgtcc agccatctgt 6840  
tgtttgcccc tccccgtgc cttccttgac cctggaaggt gccactcca ctgtcctttc 6900  
ctaataaaat gaggaattg catcgcatg tctgagtagg tgatattcta ttctgggggg 6960  
tgggggtggg cagcacagca agggggagga ttgggaagac aatagcaggc atgctgggga 7020  
tgccgtgggc tctatgggta cccagggtgt gaagaattga cccggttcct cctgggccag 7080  
aaagaagcag gcacatcccc ttctctgtga cacaccctgt ccacgcccct ggttcttagt 7140  
tccagcccca ctcataggac actcatagct caggagggtt ccgccttcaa tcccaccgc 7200  
taaagtactt ggagcgggtt ctccctccct catcagccca ccaaaccaaa cctagcctcc 7260  
aagagtggga agaaattaaa gcaagatagg ctattaagt cagagggaga gaaaatgcct 7320  
ccaacatgtg aggaagtaat gagagaaatc atagaatttc ttccgcttcc tcgctcactg 7380  
actcgtgcg ctccgtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 7440  
tacgggttat cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 7500  
aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt ttcccatagg ctccgcccc 7560  
ctgacgagca tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat 7620  
aaagatacca ggcgtttccc cctggaagct ccctcgtgc ctctcctgtt ccgaccctgc 7680  
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct 7740  
cacgctgtag gtatctcagt tcgggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 7800  
aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 7860  
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 7920  
ggtatgtagg cggtgctaca gatttcttga agtggtggcc taactacggc tacactagaa 7980  
ggacagtatt tggatatctg gctctgctga agccagttac cttcggaaaa agagtgtgta 8040  
gctcttgatc cggcaaaaa accaccgctg gtagcgggtg tttttttgtt tgcaagcagc 8100  
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg 8160  
acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 8220  
tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 8280

```

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 8340
gtctattttcg ttcattccata gttgcctgac tccggggggg gggggcgctg aggtctgcct 8400
cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa 8460
gtgagggagc cacggttgat gagagctttg ttgtaggtgg accagttggg gattttgaac 8520
ttttgctttg ccacggaacg gtctgcgttg tgggaagat gcgtgatctg atccttcaac 8580
tcagcaaaaag ttcgatttat tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct 8640
gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa 8700
actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta 8760
atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcgggtctg 8820
cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt 8880
tatcaagtga gaaatcacca tgagtgcga ctgaatccgg tgagaatggc aaaagcttat 8940
gcattttctt ccagacttgt tcaacaggcc agccattacg ctgctcatca aaatcactcg 9000
catcaaccaa accgttatct attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc 9060
tggtaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaa actgccagcg 9120
catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc 9180
cggggatcgc agtgggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg 9240
tcggaagagg cataaattcc gtcagccagt ttagtctgac catctcatct gtaacatcat 9300
tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcctcgggc ttcccataca 9360
atcgatagat tgcgcacct gattgcccgga cattatcgcg agcccattta taccatata 9420
aatcagcatc catgttgga tttaatcgcg gcctcgagca agacgtttcc cggtgaatat 9480
ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg 9540
atatattttt atcttggtgca atgtaacatc agagattttg agacacaacg tggctttccc 9600
cccccccca ttattgaagc atttatcagg gttattgtct catgagcgga tacatatctg 9660
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 9720
ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga 9780
ggccctttcg tc 9792

```

&lt;210&gt; 171

&lt;211&gt; 9167

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 171

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggg ttgggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggatgcattg gaacgcggat 1020

```



tccccgtgcc	aagagtgcgc	taagtaccgc	ctatagactc	tataggcaca	cccctttggc	1080
tcttatgcat	gctatactgt	ttttggcttg	gggcctatac	acccccgctt	ccttatgcta	1140
taggtgatgg	tatagcttag	cctataggtg	tgggttattg	accattattg	accactcccc	1200
tattggtgac	gatactttcc	attactaatc	cataacatgg	ctctttgcca	caactatctc	1260
tattggctat	atgccaatac	tctgtccttc	agagactgac	acggactctg	tatttttaca	1320
ggatggggtc	ccattttatta	tttaciaaatt	cacatatata	acaacgcctg	cccccggtgc	1380
cgcagttttt	attaaacata	gcgtggggatc	tccacgcgaa	tctcgggtac	gtgttccgga	1440
catgggctct	tctccggtag	cggcggagct	tccacatccg	agccctgggtc	ccatgcctcc	1500
agcggctcat	ggtcgctcgg	cagctccttg	ctcctaacag	tggaggccag	acttaggcac	1560
agcacaatgc	ccaccaccac	cagtgtgccg	cacaaggccg	tggcggtagg	gtatgtgtct	1620
gaaaatgagc	gtggagattg	ggctcgcacg	gctgacgcag	atggaagact	taaggcagcg	1680
gcagaagaag	atgcaggcag	ctgagttgtt	gtattctgat	aagagtcaga	ggtaactccc	1740
gttgcggtgc	tgttaacggg	ggagggcagt	gtagtctgag	cagtactcgt	tgctgccgcg	1800
cgcgccacca	gacataatag	ctgacagact	aacagactgt	tcctttccat	gggtcttttc	1860
tgcagtcacc	gtcgtcgaca	tggcccgcgc	cagcgtgctg	agcggcggcg	agctggaccg	1920
ctgggagaag	atccgcctgc	gccccggcgg	caagaagaag	tacaagctga	agcacatcgt	1980
gtgggccagc	cgcgagctgg	agcgtctcgc	cgtgaacccc	ggcctgctgg	agaccagcga	2040
gggctgccgc	cagatcctgg	gccagctgca	gcccagcctg	cagaccggca	gcgaggagct	2100
gcgcagcctg	tacaacaccg	tggccaccct	gtactgcgtg	caccagcgca	tcgagatcaa	2160
ggacaccaag	gaggccctgg	acaagatcga	ggaggagcag	aacaagagca	agaagaaggc	2220
ccagcaggcc	gccgccgaca	ccggccacag	caaccagggtg	agccagaact	accccatcgt	2280
gcagaacatc	cagggccaga	tgggtgcacca	ggccatcagc	ccccgcaccc	tgaacgcctg	2340
ggtgaagggtg	gtggaggaga	aggccttcag	ccccgagggtg	atccccatgt	tcagcgccct	2400
gagcgagggc	gccaccccc	aggacctgaa	caccatgctg	aacaccgtgg	gcggccacca	2460
ggccgccatg	cagatgctga	aggagaccat	caacgaggag	gccgccgagt	gggaccgcgt	2520
gcaccccggtg	cacgccggcc	ccatcgcccc	cggccagatg	cgcgagcccc	gcggcagcga	2580
catcgccggc	accaccagca	ccctgcagga	gcagatcggc	tggatgacca	acaaccccc	2640
catccccgtg	ggcgagatct	acaagcgctg	gatcatcctg	ggcctgaaca	agatcggtgcg	2700
catgtacagc	cccaccagca	tcctggacat	ccgccagggc	cccaaggagc	ccttcgcgca	2760
ctacgtggac	cgcttctaca	agaccctgcg	cgccgagcag	gccagccagg	aggtgaagaa	2820
ctggatgacc	gagaccctgc	tgggtgcagaa	cgccaacccc	gactgcaaga	ccatcctgaa	2880
ggccctgggc	cccgccgcca	ccctggagga	gatgatgacc	gcctgccagg	gcgtgggcgg	2940
ccccggccac	aaggcccgcg	tgctggccga	ggccatgagc	caggtgacca	acagcgccac	3000
catcatgatg	cagcgcggca	acttccgcaa	ccagcgcaag	atcgtgaagt	gcttcaactg	3060
cggcaaggag	ggccacaccg	cccgcactg	ccgcgcccc	cgcaagaagg	gctgctggaa	3120
gtgcggaag	gagggccacc	agatgaagga	ctgcaccgag	cgacaggcta	ataggggaaga	3180
tctggccttc	ccacaaggga	aggccaggga	atcttcttca	gagcagacca	gagccaacag	3240
ccccaccaga	agagagcttc	aggtttgggg	aagagacaac	aactccctct	cagaagcagg	3300
agccgataga	caagggaactg	tatcctttag	cttccctcag	atcactcttt	ggcagcgacc	3360
cctcgtcaca	ataaagatag	ggggccagct	gaaggaggcc	cttctagaca	ccggcgccga	3420
cgacaccgtg	ctggaggaga	tgaacctgcc	cggccgctgg	aagcccaaga	tgatcggcgg	3480
catcggcggc	ttcatcaagg	tgggccagta	cgaccagatc	ctgatcgaga	tctgcgccca	3540
caaggccatc	ggcaccgtgc	tgggtggccc	cacccccgtg	aacatcatcg	gccgcaacct	3600
gctgaccag	atcggtgca	ccctgaactt	ccccatcagc	cccatcgaga	ccgtgcccg	3660
gaagctgaag	cccggcatgg	acggcccca	ggtgaagcag	tggccctga	ccgaggagaa	3720
gatcaaggcc	ctggtggaga	tctgcaccga	gatggagaag	gagggcaaga	tcagcaagat	3780
cggccccgag	aacccttaca	acacccccgt	gttcgccatc	aagaagaagg	acagcaccaa	3840
gtggcgcaag	ctggtggact	tccgcgagct	gaacaagcgc	acccaggact	tctgggaggt	3900
gcagctgggc	atccccacc	ccgcggcct	gaagcagaag	aagagcgtga	ccgtgctgga	3960
cgtgggcgac	gcctacttca	gcgtgcccct	ggacaaggac	ttccgcaagt	acaccgcctt	4020
caccatcccc	agcatcaaca	acgagacccc	cggcatccgc	taccagtaca	acgtgctgcc	4080
ccagggtcgg	aagggcagcc	ccgccatctt	ccagtgcagc	atgaccaaga	tcctggagcc	4140
cttccgcaag	cagaaccccc	acatcgtgat	ctaccagtac	atggaccacc	tgtacgtggg	4200
cagcgacctg	gagatcgccc	agcaccgcac	caagatcgag	gagctgcgcc	agcacctgct	4260
gcgctggggc	ttcaccaccc	ccgacaagaa	gcaccagaag	gagccccctt	tcctgtggat	4320

```

gggctacgag ctgcaccccg acaagtggac cgtgcagccc atcgtgctgc ccgagaagga 4380
cagctggacc gtgaacgaca tccagaagct ggtgggcaag ctgaactggg ccagccagat 4440
ctacgccggc atcaagggtgc gccagctgtg caagctgctg cgccggcacca aggccctgac 4500
cgaggtggtg cccctgaccg aggaggccga gctggagctg gccgagaacc gcgagatcct 4560
gaaggagccc gtgcacggcg tgtactacga cccagcaag gacctgatcg ccgagatcca 4620
gaagcagggc cagggccagt ggacctacca gatctaccag gagcccttca agaacctgaa 4680
gaccggcaag tacgcccgca tgaagggcgc ccacaccaac gacgtgaagc agctgaccga 4740
ggccgtgcag aagatcgcca ccgagagcat cgtgatctgg ggcaagaccc ccaagttcaa 4800
gctgcccac cagaaggaga cctgggaggg ctggtggacc gagtactggc aggccacctg 4860
gatccccgag tgggagttcg tgaacacccc cccctgggtg aagctgtggt accagctgga 4920
gaaggagccc atcatcggcg ccgagacctt ctacgtggac ggcgcgcgca accgcgagac 4980
caagctgggc aaggccggct acgtgaccga ccgcggccgc cagaagggtg tgcccctgac 5040
cgacaccacc aaccagaaga ccgagctgca ggccatccac ctggccctgc aggacagcgg 5100
cctggaggtg aacatcgtga ccgacagcca gtacgccctg ggcacatcc aggccagacc 5160
cgacaagagc gagagcgagc tgggtgagcca gatcatcgag cagctgatca agaaggagaa 5220
ggtgtacctg gcctgggtgc ccgccacaa gggcatcggc ggcaacgagc aggtggacgg 5280
cctggtgagc gccggcatcc gcaagggtgct gttcctggac ggcacgcaca aggccagga 5340
ggagcacgag aagtaccaca gcaactggcg cgccatggcc agcgacttca acctgcccc 5400
cgtggtggcc aaggagatcg tggccagctg cgacaagtgc cagctgaagg gcgaggccat 5460
gcacggccag gtggactgca gcccggcat ctggcagctg gcatgcacc acctggaggg 5520
caaggtgatc ctggtggccg tgcacgtggc cagcggctac atcgaggccg aggtgatccc 5580
cgccgagacc ggccaggaga ccgcctactt cctgctgaag ctggccggcc gctggcccgt 5640
gaagaccgtg cacaccgaca acggcagcaa cttcaccagc accaccgtga aggccgcctg 5700
ctggtgggccc ggcacaaagc aggagtccgg catcccctac aacccccaga gccaggcgct 5760
gatcgagagc atgaacaagg agctgaagaa gatcatcggc caggtgcgcg accaggccga 5820
gcacctgaag accgcctgac agatggcctg gttcatccac aacttcaagc gcaaggcgcg 5880
catcggcggc tacagcgccg gcgagcgcat cgtggacatc atcgccaccg acatccagac 5940
caaggagctg cagaagcaga tcaccaagat ccagaacttc cgcgtgtact accgcgacag 6000
ccgcgacccc gtgtggaagg gcccgcgcaa gctgctgtgg aagggcgagg gcgcctggt 6060
gatccaggac aacagcgaca tcaagggtgg gcccgcgcg aaggccaaga tcatccgca 6120
ctacggcaag cagatggccg gcgacgactg cgtggccagc cgccaggacg aggactagga 6180
attcttgctg tgccttctag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc 6240
ttgaccctgg aaggtgccac tcccactgtc ctttctctaat aaaatgagga aattgcatcg 6300
cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagca cagcaagggg 6360
gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat gggtagccag 6420
gtgctgaaga attgaccggg ttctctctgg gccagaaaga agcaggcaca tccccttctc 6480
tgtgacacac cctgtccacg cccctggttc ttagttccag cccactcat aggacactca 6540
tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttggagc ggtctctccc 6600
tccctcatca gccacccaaa ccaaacctag cctccaagag tgggaagaaa ttaaagcaag 6660
ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 6720
aaatcataga atttcttccg ctctctcgct cactgactcg ctgcgctcgg tcgttcggct 6780
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 6840
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 6900
cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 6960
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 7020
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 7080
tctcccttcg ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcggg 7140
gtaggtcggt cgctccaagc tgggctgtgt gcacgaacc cccgttcagc ccgaccgctg 7200
cgcttatcc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact 7260
ggcagcagcc actggtaaca ggattagcag agcgaggat gtaggcggtg ctacagagtt 7320
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatbtggta tctgcgctct 7380
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 7440
cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7500
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg 7560
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 7620

```

```

aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttggtctg acagttacca 7680
atgcttaatc agtgaggcac ctatctcagc gatctgtcta ttctgttcat ccatagttgc 7740
ctgactccgg gggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat 7800
accaggcctg aatcgcccca tcatccagcc agaaagtgag ggagccacgg ttgatgagag 7860
ctttgttgta ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggtctg 7920
cgttgtcggg aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac 7980
aaagccgccc tcccgtaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa 8040
ttctgattag aaaaactcat cgagcatcaa atgaaactgc aattttattca tatcaggatt 8100
atcaatacca tttttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca 8160
gttccatagg atggcaagat cctgggtatcg gtctgcgatt ccgactcgtc caacatcaat 8220
acaacctatt aattttccct cgtcaaaaat aagggttatca agtgagaaat caccatgagt 8280
gacgactgaa tccggtgaga atggcaaaaag cttatgcatt tctttccaga cttgttcaac 8340
aggccagcca ttacgctcgt catcaaaaatc actcgcatca accaaaccgt tattcattcg 8400
tgattgcgcc tgagcgagac gaaatacgcg atcgtctgta aaaggacaat tacaacagg 8460
aatcgaatgc aaccggcgca ggaacactgc cagcgcatca acaatatttt cacctgaatc 8520
aggatattct tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca 8580
tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag 8640
ccagtttagt ctgaccatct catctgtaac atcattggca acgtacctt tgccatgttt 8700
cagaaacaac tctggcgcgt cgggcttccc atacaatcga tagattgtcg cacctgattg 8760
cccagacatta tcgcgagccc atttataccc atataaatca gcatccatgt tggaaattta 8820
tcgcggcctc gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact 8880
gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 8940
acatcagaga ttttgagaca caacgtggct tcccccccc cccattatt gaagcattta 9000
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 9060
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 9120
catgacatta acctataaaa ataggcggtat caccgagccc tttcgtc 9167

```

&lt;210&gt; 172

&lt;211&gt; 9783

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 172

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttgggtcatg 300
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
ccgcctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020

```



```

tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgtt ctttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatatata acaacgccgt cccccgtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgcagtcacc gtcgtcgaca tggcccgcgc cagcgtgctg agcggcggcg agctggaccg 1920
ctgggagaag atccgcctgc gccccggcgg caagaagaag tacaagctga agcacatcgt 1980
gtggggcagc cgcgagctgg agcgttcgc cgtgaacccc ggctgtctgg agaccagcga 2040
gggctgccgc cagatcctgg gccagctgca gccagcctg cagaccggca gcgaggagct 2100
gcgcagcctg tacaacaccg tggccaccct gtactgctg caccagcga tcgagatcaa 2160
ggacaccaag gaggccctgg acaagatcga ggaggagcag aacaagagca agaagaaggc 2220
ccagcaggcc gccgccgaca ccggccacag caaccaggtg agccagaact accccatcgt 2280
gcagaacatc cagggccaga tgggtgcacca ggccatcagc ccccgcaacc tgaacgcctg 2340
ggtgaagggt gtggaggaga aggccttcag ccccgagggt atccccatgt tcagcgccct 2400
gagcgagggc gccaccccc aggacctgaa caccatgctg aacaccgtgg gcggccacca 2460
ggccgccatg cagatgctga aggagaccat caacgaggag gccgccgagt gggaccgcgt 2520
gcaccccggt cagcggccgc ccatcgcccc cggccagatg cgcgagcccc gcggcagcga 2580
catcgccggc accaccagca ccctgcagga gcagatcggc tggatgacca acaaccccc 2640
catccccgtg ggcgagatct acaagcgtg gatcatcctg ggcctgaaca agatcgtgcg 2700
catgtacagc cccaccagca tcctggacat ccgccagggc cccaaggagc ccttccgcga 2760
ctacgtggac cgcttctaca agaccctgcg cgcgagcag gccagccagg aggtgaagaa 2820
ctggatgacc gagaccctgc tgggtgcagaa cgccaacccc gactgcaaga ccatcctgaa 2880
ggccctgggc cccgccgcca ccctggagga gatgatgacc gcctgccagg gcgtgggcgg 2940
ccccggccac aaggcccgcg tgcgtggcga ggccatgagc caggtgacca acagcgccac 3000
catcatgatg cagcgcggca acttcgcga ccagcgcaag atcgtgaagt gcttcaactg 3060
cggcaaggag ggccacaccg cccgcaactg ccgcgcccc cgcaagaagg gctgctggaa 3120
gtgcggcaag gaggccacc agatgaagga ctgcaccgag cgacaggcta atagggaaga 3180
tctggccttc ccacaaggga aggccaggga attttcttca gagcagacca gagccaacag 3240
ccccaccaga agagagcttc aggtttgggg aagagacaac aactccctct cagaagcagg 3300
agccgataga caaggaaactg tacccttag ctccctcag atcactcttt ggcagcgacc 3360
cctcgtcaca ataaagatag ggggcccagct gaaggaggcc cttctagaca ccggcgccga 3420
cgacaccgtg ctggaggaga tgaacctgcc cggccgctgg aagcccaaga tgatcggcgg 3480
catcggcggc ttcatacaagg tgggcccagta cgaccagatc ctgatcgaga tctgcggcca 3540
caaggccatc ggcaccgtgc tgggtggccc cacccccgtg aacatcatcg gccgcaacct 3600
gctgaccagc atcggtgca ccctgaactt ccccatcagc cccatcgaga ccgtgccctg 3660
gaagctgaag cccggcatgg acggcccca ggtgaagcag tggcccttga ccgaggagaa 3720
gatcaaggcc ctggtggaga tctgcaccga gatggagaag gagggaaga tcagcaagat 3780
cggccccgag aacccttaca acaccccggt gtccgccatc aagaagaagg acagcaccaa 3840
gtggcgcaag ctggtggact tccgcgagct gaacaagcgc acccaggact tctgggaggt 3900
gcagctgggc atccccacc ccgcccgcct gaagcagaag aagagcgtga ccgtgctgga 3960
cgtggggcgc gcctacttca gcgtgccct ggacaaggac ttcgcgaagt acaccgcctt 4020
caccatcccc agcatcaaca acgagacccc cggcatccgc taccagtaca acgtgctgcc 4080
ccagggtggg aagggcagcc ccgccatctt ccagtgcagc atgaccaaga tcctggagcc 4140
cttccgcaag cagaacccc acatcgtgat ctaccagtac atggaccacc tgtacgtggg 4200
cagcgacctg gagatcggcc agcaccgcac caagatcgag gagctgcgcc agcacctgct 4260
gcgctggggc ttcaccaccc ccgacaagaa gcaccagaag gagccccct tcctgtggat 4320

```

```

gggctacgag ctgcacccccg acaagtggac cgtgcagccc atcgtgctgc ccgagaagga 4380
cagctggacc gtgaacgaca tccagaagct ggtgggcaag ctgaactggg ccagccagat 4440
ctacgccggc atcaaggtgc gccagctgtg caagctgctg cgcggcacca aggccctgac 4500
cgaggtggtg cccctgaccg aggaggccga gctggagctg gccgagaacc gcgagatcct 4560
gaaggagccc gtgcacggcg tgtactacga ccccagcaag gacctgatcg ccgagatcca 4620
gaagcagggc cagggccagt ggacctacca gatctaccag gagcccttca agaacctgaa 4680
gaccggcaag tacgcccgcg tgaagggcgc ccacaccaac gacgtgaagc agctgaccga 4740
ggcgtgcag aagatcgcca ccgagagcat cgtgatctgg ggcaagaccc ccaagttcaa 4800
gctgcccac cagaaggaga cctgggaggg ctggtggacc gactactggc aggccacctg 4860
gatccccgag tgggagttcg tgaacacccc cccctgggtg aagctgtggt accagctgga 4920
gaaggagccc atcatcggcg ccgagacctt ctacgtggac ggcggccgca accgcgagac 4980
caagctgggc aaggccggct acgtgaccga ccgcggccgc cagaaggtgg tggccctgac 5040
cgacaccacc aaccagaaga ccgagctgca ggccatccac ctggccctgc aggacagcgg 5100
cctggaggtg aacatcgtga ccgacagcca gtacgccctg ggcatcatcc aggccagacc 5160
cgacaagagc gagagcgagc tgggtgagcca gatcatcgag cagctgatca agaaggagaa 5220
ggtgtacctg gcctgggtgc ccgccacaa gggcatcggc ggcaacgagc aggtggacgg 5280
cctggtgagc gccggcatcc gcaaggtgct gtccctggac ggcatcgaca aggccagga 5340
ggagcacgag aagtaccaca gcaactggcg cgccatggcc agcgacttca acctgcccc 5400
cgtggtggcc aaggagatcg tggccagctg cgacaagtgc cagctgaagg gcgaggccat 5460
gcacggccag gtggactgca gccccggcat ctggcagctg gcatgcaccc acctggaggg 5520
caaggtgatc ctggtggccg tgcacgtggc cagcggctac atcgaggccg aggtgatccc 5580
cgccgagacc ggccaggaga ccgcctactt cctgctgaag ctggccggcc gctggcccg 5640
gaagaccgtg cacaccgaca accgcagcaa cttaccagc accaccgtga aggccgctg 5700
ctggtgggccc ggcatcaagc aggagtccg catccccctac aacccccaga gccagggcg 5760
gatcgagagc atgaacaagg agctgaagaa gatcatcggc caggtgcgcg accaggccga 5820
gcacctgaag accgccgtgc agatggccgt gtccatccac aacttcaagc gcaagggcg 5880
catcgccggc tacagcgccg gcgagcgcat cgtggacatc atcgccaccg acatccagac 5940
caaggagctg cagaagcaga tcaccaagat ccagaacttc cgcgtgtact accgcgacag 6000
ccgcgacccc gtgtggaagg gccccgcaa gctgctgtgg aagggcgagg gcgccgtgg 6060
gatccaggac aacagcgaca tcaaggtggt gccccgcgc aaggccaaga tcatccgca 6120
ctacggcaag cagatggccg gcgacgactg cgtggccagc cgccaggacg aggacatggg 6180
cggcaagtgg agcaagagca gcgtgatcgg ctggcccgc gtgcgcgagc gcatgcgccg 6240
cgccgagccc gccgcccagc gcgtgggccc cgtgagccgc gacctggaga agcacggcgc 6300
catcaccagc agcaacaccg ccgccaacaa cgccgcctgc gcctggctgg aggccagga 6360
ggaggaggag gtgggcttcc ccgtgacccc ccaggtgccc ctgcgccccca tgacctaaa 6420
ggccgcccgtg gacctgagcc acttccctgaa ggagaagggc ggccctggagg gcctgatcca 6480
cagccagcgc cgccaggaca tcctggacct gtggatctac cacaccagc gctacttccc 6540
cgactggcag aactacaccc ccggccccgg cgtgcgctac cccctgacct tcggctggtg 6600
ctacaagctg gtgcccgtgg agcccgacaa ggtggaggag gccaacaagg gcgagaacac 6660
cagcctgctg caccgccgtg gcctgcacgg catggacgac cccgagcgcg aggtgctgga 6720
gtggcgcttc gacagccgcc tggccttcca ccacgtggcc cgcgagctgc accccgagta 6780
cttcaagaac tgctgaggat caattcttgc tgtgccttct agttgccagc catctgttgt 6840
ttgcccctcc ccgctgcctt ccttgacctt ggaaggtgcc actcccactg tcctttccta 6900
ataaaatgag gaaattgcat cgcattgtct gagtagtgtt cattctatct tggggggtgg 6960
ggtggggcag cacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 7020
ggtgggctct atgggtaccc aggtgctgaa gaattgaccc ggttcctcct gggccagaaa 7080
gaagcaggca catccccttc tctgtgacac accctgtcca cgcccctggg tcttagttcc 7140
agccccactc ataggacact catagctcag gagggctccg ccttcaatcc caccgctaa 7200
agtacttgga gcggtctctc cctccctcat cagcccacca aaccaaacct agcctccaag 7260
agtgggaaga aattaaagca agataggcta ttaagtgcag agggagagaa aatgcctcca 7320
acatgtgagg aagtaatgag agaaatcata gaatttcttc cgcttccctg ctcactgact 7380
cgctgcctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatat 7440
ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 7500
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 7560
acgagcatca caaaaatcga cgctcaagtc agagggtggc aaacccgaca ggactataaa 7620

```



```

gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 7680
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgtttctt caatgctcac 7740
gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 7800
cccccgttca gcccgaaccg tgcgccttat ccggttaacta tcgtcttgag tccaaccgg 7860
taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 7920
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 7980
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 8040
cttgatccgg caaacaacc accgctggta gcggtggtgt ttgcaagcag cagattacgc 8100
gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacgggggtct gacgctcagt 8160
ggaacgaaaa ctcacgttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct 8220
agatcctttt aaattaaaaa tgaagtttta aatcaatcta agtatatat gagtaacctt 8280
ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc 8340
gttcatccat agttgcctga ctccgggggg ggggggcgct gaggtctgcc tcgtgaagaa 8400
ggtgttgctg actcatacca ggcctgaatc gccccatcat ccagccagaa agtgaggagg 8460
ccacggttga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa cttttgcttt 8520
gccacggaac ggtctgcgct gtccgggaaga tgcgtgatct gatccttcaa ctcagcaaaa 8580
gttcgattta ttcaacaag ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt 8640
acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt 8700
tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag 8760
aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga 8820
ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagt 8880
agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagctta tgcatttctt 8940
tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc gcatcaacca 9000
aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag 9060
gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa 9120
tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg 9180
cagtggtagg taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag 9240
gcataaattc cgtcagccag tttagtctga ccattctcatc tgtaacatca ttggcaacgc 9300
tacctttgcc atgtttcaga aacaactctg gcgcacggg cttcccatc aatcgataga 9360
ttgtcgcacc tgattgcccg acattatcgc gagcccatth ataccatata aaatcagcat 9420
ccatgttggg atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa 9480
cacccttgtt attactgttt atgtaagcag acagttttat tgttcatgat gatataattt 9540
tatcttgtgc aatgtaacat cagagatttt gagacacaa gtggctttcc ccccccccc 9600
attattgaag catttatcag ggttattgtc tcatgagcgg atacataatt gaatgtattt 9660
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct 9720
aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 9780
gtc

```

&lt;210&gt; 173

&lt;211&gt; 9782

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 173

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360

```

```

ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattgg 420
cccgccctggc tgaccgcccc acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccccgct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggagggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgccctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcgc taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattgggtgac gatactttcc attactaatc cataacatgg ctcttttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgcctg ccccgctgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttcggga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860
tgcatgcacc gtcgtcgaca tggcaaatg gagcaagagc agcgtgatcg gctggccccg 1920
cgtgcgcgag cgcattgcgc gcgccgagcc cgcgcgcgac ggcgtgggag cgtgagccg 1980
cgacctggag aagcacggcg ccacaccag cagcaacacc gccgccaaca acgcccctg 2040
cgcctggctg gagggccagg aggaggagga ggtgggcttc cccgtgacct cccaggtgcc 2100
cctgcgcccc atgacctaca aggcgcgct ggacctgagc cacttcctga aggagaagg 2160
cggcctggag ggctgatcc acagccagcg ccgccaggac atcctggacc tgtggatcta 2220
ccacaccag ggctacttcc ccgactggca gaactacacc cccggccccg gcgtgcgcta 2280
ccccctgacc ttcggctggt gctacaagct ggtgcccgtg gagcccgaca aggtggagga 2340
ggccaacaag ggcgagaaca ccagcctgct gcaccccgtg agcctgcacg gcatggacga 2400
ccccgagcgc gaggtgctgg agtggcgctt cgacagccgc ctggccttcc accacgtggc 2460
ccgcgagctg caccgcgagt acttcaagaa ctgcatggcc cgcgccagcg tgctgagcgg 2520
cggcgagctg gaccgctggg agaagatccg cctgcgcccc ggcggcaaga agaagtacaa 2580
gctgaagcac atcgtgtggg ccagccgcga gctggagcgc ttcgccgtga accccggcct 2640
gctggagacc agcgagggtt gccgccagat cctgggccag ctgcagccca gcctgcagac 2700
cggcagcgag gagctgcgca gcctgtacaa caccgtggcc accctgtact gcgtgcacca 2760
gcgcatcgag atcaaggaca ccaaggaggc cctggacaag atcgaggagg agcagaacaa 2820
gagcaagaag aaggcccagc aggcgcgcgc cgacaccggc cacagcaacc aggtgagcca 2880
gaactacccc atcgtgcaga acatccaggg ccagatgggt caccaggcca tcagcccccg 2940
caccctgaac gcctgggtga aggtgggtgga ggagaaggcc ttcagccccg aggtgatccc 3000
catgttcagc gccctgagcg agggcgccac ccccaggac ctgaacacca tgctgaacac 3060
cgtgggcggc caccaggccg ccatgcagat gctgaaggag accatcaacg aggaggccgc 3120
cgagtgggac cgcgtgcacc ccgtgcacgc cggccccatc gccccggcc agatgcgcga 3180
gccccgcggc agcgacatcg ccggcaccac cagcaccctg caggagcaga tcggctggat 3240
gaccaacaac cccccatcc ccgtgggcca gatctacaag cgctggatca tctgggcct 3300
gaacaagatc gtgcgcatgt acagccccac cagcatcctg gacatccgcc agggcccca 3360
ggagcccttc cgcgactacg tggaccgctt ctacaagacc ctgcgcgcc agcaggccag 3420
ccaggaggtg aagaactgga tgaccgagac cctgctgggt cagaacgcca accccgactg 3480
caagaccatc ctgaaggccc tgggccccgc cgcaccctg gaggagatga tgaccgctg 3540
ccagggcgtg ggcggccccg gccacaaggc ccgcgtgctg gccgaggcca tgagccaggt 3600
gaccaacagc gccaccatca tgatgcagcg cggcaacttc cgcaaccagc gcaagatcgt 3660

```

gaagtgcttc aactgcggca aggagggcca caccgcccgc aactgccgcg cccccgcaa 3720  
gaagggctgc tggaagtgcg gcaaggaggg ccaccagatg aaggactgca ccgagcgaca 3780  
ggctaatagg gaagatctgg ccttcccaca agggaaggcc agggaaatttt cttcagagca 3840  
gaccagagcc aacagcccca ccagaagaga gcttcagggt tggggaagag acaacaactc 3900  
cctctcagaa gcaggagccg atagacaagg aactgtatcc tttagcttcc ctcagatcac 3960  
tctttggcag cgacccctcg tcacaataaa gatagggggc cagctgaagg aggccttct 4020  
agacaccggc gccgacgaca ccgtgctgga ggagatgaac ctgcccggcc gctggaagcc 4080  
caagatgacg ggcggcatcg gcggcttcat caaggtgggc cagtacgacc agatcctgat 4140  
cgagatctgc ggccacaagg ccatcggcac cgtgctggtg ggccccaccc ccgtgaacat 4200  
catcgccgcg aacctgctga ccagatcgg ctgcacctg aacttcccca tcagcccat 4260  
cgagaccgtg cccgtgaagc tgaagcccg catggacggc cccaagggtg agcagtggcc 4320  
cctgaccgag gagaagatca aggccttgg ggagatctgc accgagatgg agaaggagg 4380  
caagatcagc aagatcggcc ccgagaacc ctacaacacc cccgtgttcg ccatcaagaa 4440  
gaaggacagc accaagtggc gcaagctggt ggacttccgc gagctgaaca agcgcacca 4500  
ggacttctgg gaggtgcagc tgggcatccc ccaccccgcc ggcctgaagc agaagaagag 4560  
cgtgaccgtg ctggacgtgg gcgacgccta cttcagcgtg cccctggaca aggacttccg 4620  
caagtacacc gccttcacca tcccagcat caacaacgag acccccgcca tccgctacca 4680  
gtacaacgtg ctgccccagg gctggaaggg cagccccgc atcttccagt gcagcatgac 4740  
caagatcctg gagcccttcc gcaagcagaa ccccgacatc gtgatctacc agtacatgga 4800  
ccacctgtac gtgggcagcg acctggagat cggccagcac cgcaccaaga tcgaggagct 4860  
gcgccagcac ctgctgcgtt ggggcttcac cacccccgac aagaagcacc agaaggagcc 4920  
ccccttcttg tggatgggct acgagctgca ccccgacaag tggaccgtgc agcccatcgt 4980  
gctgcccag aaggacagct ggaccgtgaa cgacatccag aagctggtgg gcaagctgaa 5040  
ctgggccagc cagatctacg ccggcatcaa ggtgcgccag ctgtgcaagc tgctgcgcgg 5100  
caccaaggcc ctgaccgagg tgggtgccct gaccgaggag gccgagctgg agctggccga 5160  
gaaccgcgag atcctgaagg agcccggtga cggcgtgtac tacgaccca gcaaggacct 5220  
gatcgccgag atccagaagc agggccaggg ccagtggacc taccagatct accaggagcc 5280  
cttcaagaac ctgaagaccg gcaagtacgc ccgcatgaag ggcgcccaca ccaacgacgt 5340  
gaagcagctg accgaggccg tgcagaagat cgccaccgag agcatcgtga tctggggcaa 5400  
gacccccaag ttcaagctgc ccatccagaa ggagacctg gaggcctgg ggaccgagta 5460  
ctggcaggcc acctggatcc ccgagtggga gtctgtgaac accccccccc tgggtgaagct 5520  
gtggtaccag ctggagaagg agcccatcat cggcgccgag accttctacg tggacggcgc 5580  
cgccaaccgc gagaccaagc tgggcaaggc cggctacgtg accgaccgcg gccgccagaa 5640  
ggtggtgccc ctgaccgaca ccaccaacca gaagaccgag ctgcaggcca tccacctggc 5700  
cctgcaggac agcggcctgg aggtgaacat cgtgaccgac agccagtacg ccctgggcat 5760  
catccaggcc cagcccgaca agagcgagag cgagctggtg agccagatca tcgagcagct 5820  
gatcaagaag gagaagggtg acctggcctg ggtgcccgc cacaaggga tcggcggcaa 5880  
cgagcagggt gacggcctgg tgagcgccgg catccgcaag gtgctgttcc tggacggcat 5940  
cgacaaggcc caggaggagc acgagaagta ccacagcaac tggcgcgcca tggccagcga 6000  
cttcaacctg ccccccgtgg tggccaagga gatcgtggcc agctgcgaca agtgccagct 6060  
gaagggcgag gccatgcacg gccagggtgga ctgcagcccc ggcactctggc agctggcatg 6120  
caccacctg gagggcaagg tgatcctgg ggccgtgcac gtggccagcg gctacatcga 6180  
ggccgagggt atccccgcc agaccggcca ggagaccgcc tacttctctg tgaagctggc 6240  
cggccgctgg cccgtgaaga ccgtgcacac cgacaacggc agcaacttca ccagcaccac 6300  
cgtgaaggcc gcctgctggt gggccggcat caagcaggag ttcggcatcc cctacaacc 6360  
ccagagccag ggcgtgatcg agagcatgaa caaggagctg aagaagatca tcggccagg 6420  
gcgcgaccag gccgagcacc tgaagaccgc cgtgcagatg gccgtgttca tccacaactt 6480  
caagcgcaag ggcggcatcg gcggctacag cgccggcgag cgcactctgg acatcatcgc 6540  
caccgacatc cagaccaagg agctgcagaa gcagatcacc aagatccaga acttccgcgt 6600  
gtactaccgc gacagccgcg acccgtgtg gaagggcccc gccaaagtgc tgtggaagg 6660  
cgagggcgcc gtggtgatcc aggacaacag cgacatcaag gtggtgcccc gccgcaagg 6720  
caagatcatc cgcgactacg gcaagcagat ggccggcgac gactgcgtgg ccagccgcca 6780  
ggacgaggac taggaattct tgctgtgcct tctagtgtcc agccatctgt tgtttgcccc 6840  
tccccgtgc cttccttgac cctggaaggt gccactccca ctgtccttct ctaataaaat 6900  
gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 6960



```

cagcacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc 7020
tctatgggta cccaggtgct gaagaattga cccggttcct cctgggccag aaagaagcag 7080
gcacatcccc ttctctgtga cacacctgt ccacgccctt gggtcttagt tccagcccca 7140
ctcataggac actcatagct caggagggct ccgccttcaa tcccaccgcg taaagtactt 7200
ggagcgggtct ctccctccct catcagccca ccaaaccaaa cctagcctcc aagagtggga 7260
agaaattaaa gcaagatagg ctattaagt cagagggaga gaaaatgcct ccaacatgtg 7320
aggaagtaat gagagaaatc atagaatttc ttccgcttcc tcgctcactg actcgtgctg 7380
ctcggctcgtt cggctgcggc gagcgggtatc agctcactca aaggcggtaa tacggttatc 7440
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 7500
gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca 7560
tcacaaaaat cgacgctcaa gtcagaggtg gcgaaaccgg acaggactat aaagatacca 7620
ggcgtttccc cctggaagct ccctcgtgct ctctcctgtt ccgacctgc cgcttaccgg 7680
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct cacgctgtag 7740
gtatctcagt tcgggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccctg 7800
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 7860
cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 7920
cgggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt 7980
tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 8040
cggcaaaaaa accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg 8100
cagaaaaaaa ggatctcaag aagatccctt gatcttttct acgggggtctg acgctcagtg 8160
gaacgaaaaa tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcacctc 8220
gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg 8280
gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttctg 8340
ttcatccata gttgcctgac tccggggggg gggggcgctg aggtctgcct cgtgaagaag 8400
gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa gtgagggagc 8460
cacggttgat gagagctttg ttgtaggtgg accagttggt gattttgaac ttttgctttg 8520
ccacggaacg gtctgcgttg tcgggaagat gcgtgatctg atccttcaac tcagcaaaag 8580
ttcgatttat tcaacaaagc cgccgtcccc tcaagtcagc gtaatgctct gccagtgtta 8640
caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa actgcaattt 8700
attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta atgaaggaga 8760
aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcggtctg cgattccgac 8820
tcgtccaaca tcaatacaac ctattaattt cccctcgtca aaaataaggt tatcaagtga 8880
gaaatcacca tgagtgcga ctgaatccgg tgagaatggc aaaagcttat gcatttcttt 8940
ccagacttgt tcaacaggcc agccattacg ctcgatcatc aaatcactcg catcaaccaa 9000
accgttatcc attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg 9060
acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg catcaacaat 9120
attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc cggggatcgc 9180
agtgggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg tcggaagagg 9240
cataaattcc gtcagccagt ttagtctgac catctcatct gtaacatcat tggcaacgct 9300
acctttgcca tgtttcagaa acaactctgg cgcacggggc ttcccataca atcgatagat 9360
tgtcgcacct gattgcccga cattatcgcg agccatttta taccatata aatcagcatc 9420
catgttgga tttaatcgcg gcctcgagca agacgtttcc cgttgaatat ggctcataac 9480
accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg atatatTTTT 9540
atcttggtgca atgtaacatc agagattttg agacacaacg tggctttccc cccccccca 9600
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatattg aatgtattta 9660
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta 9720
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg 9780
tc

```

&lt;210&gt; 174

&lt;211&gt; 9788

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 174

```

tcgcgcggttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccggggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tggggcggtag gcgtgtacgg tgggagggtc atataagcag 900
agctcgttta gtgaaccgtc agatcgctc gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgaag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttgg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttaaaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggatc tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500
agcggctcat ggtcgtcggc cagctccttg ctccaaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggcttttcc 1860
tgcagtcacc gtcgtcgaca tgggcgcccg cgccagcatc ctgcgcggcg gcaagctgga 1920
cgcctgggag cgcatacaagc tgaagcccgg cggcaagaag cactacatga tgaagcacct 1980
ggtgtgggac agccgcgagc tggagcgctt cgccctggac cccggcctgc tggagaccag 2040
cgagggctgc aagcagatca tgaagcagct gcagcccgcg ctgcagaccg gcaccaagga 2100
gctgatcagc ctgcacaaca ccgtggccac cctgtactgc gtgcacgaga agatcgacgt 2160
gcgcgacacc aaggaggccc tggacaagat caaggaggag cagaacaaga gccagcagaa 2220
gacccagcag gccgaggccg ccgacaaggg caaggtgagc cagaactacc ccatcgtgca 2280
gaacctgcag ggccagatgg tgcaccagcc catcagcgcc cgcaccctga acgcctgggt 2340
gaagggtggtg gaggagaagg ccttcagccc cgaggtgatc cccatgttca ccgcctgag 2400
cgagggcgcc acccccagc acctgaacac catgctgaac accgtgggcg gccaccagge 2460
cgccatgcag atgctgaagg acaccatcaa cgaggaggcc gccgagtggg accgcctgca 2520
ccccgtgcac gccggccccg tggcccccg ccagatgcgc gagccccgcg gcagcgacat 2580
cgccggcacc accagcacc tgcaggagca gatcacctgg atgaccaaca accccccgt 2640
gcccggtggc gacatctaca agcgtggat catcctgggc ctgaacaaga tcgtgcgcat 2700
gtacagcccc gtgagcatcc tggacatcaa gcagggcccc aaggagccct tccgcgacta 2760
cgtggaccgc ttcttcaaga ccctgcgcgc cgagcaggcc acccaggacg tgaagaactg 2820
gatgaccgac accctgctgg tgcagaacgc caaccccgac tgcaagacca tcctgcgcgc 2880
cctgggcccc ggcgcagcc tggaggagat gatgaccgcc tgccagggcg tgggcggccc 2940
cggccacaag gcccgcgctgc tggccgaggc catgagcaag gtgaacaaca ccaacatcat 3000

```



gatgcagcgc	agcaactgca	agggccccc	gcgcaccatc	aagtgccttca	actgcggcaa	3060
ggagggccac	ctggcccgc	actgccgcgc	cccccgcaag	aagggctgct	ggaagtgcgg	3120
caaggagggc	caccaggtga	aggactgcac	cgagcgacag	gctaataagg	aagatctggc	3180
cttcccacaa	gggaaggcca	gggaattttc	ttcagagcag	accagagcca	acagccccac	3240
cagaagagag	cttcagggtt	ggggaagaga	caacaactcc	ctctcagaag	caggagccga	3300
tagacaagga	actgtatcct	ttagcttccc	tcagatcact	ctttggcagc	gacccctcgt	3360
cacaataaag	atagggggcc	agctgaagga	ggcccttcta	gacaccggcg	ccgacgacac	3420
cgtgctggag	gagatgaacc	tgcccggccg	ctggaagccc	aagatgatcg	gcggcatcgg	3480
cggcttcac	aaggtggg	agtagcagca	gaccttgatc	gagatctg	gccacaaggc	3540
catcggcacc	gtgctggtg	gccccacccc	cgtgaacatc	atcggccgca	acctgctgac	3600
ccagatcggc	tgacccctga	acttccccat	cagcccccac	gagaccgtgc	ccgtgaagct	3660
gaagcccggc	atggacggcc	ccaaggtgaa	gcagtggccc	ctgaccgagg	agaagatcaa	3720
ggccctggtg	gagatctgca	ccgagatgga	gaaggagggc	aagatcagca	agatcggccc	3780
cgagaacccc	tacaacaccc	ccgtgttcgc	catcaagaag	aaggacagca	ccaagtggcg	3840
caagctggtg	gacttccgcg	agctgaacaa	gcgcaccacg	gacttctggg	aggtgcagct	3900
gggcatcccc	caccccgccg	gcctgaagca	gaagaagagc	gtgaccgtgc	tggacgtggg	3960
cgacgcctac	ttcagcgtgc	ccctggacaa	ggacttccgc	aagtacaccg	ccttcaccat	4020
ccccagcatc	aacaacgaga	cccccgccat	ccgctaccag	tacaacgtgc	tgccccaggg	4080
ctggaagggc	agccccgcca	tcttccagtg	cagcatgacc	aagatcctgg	agcccttccg	4140
caagcagaac	cccagacatc	tgatctacca	gtacatggac	cacctgtacg	tgggcagcga	4200
cctggagatc	ggccagcacc	gcaccaagat	cgaggagctg	cgcagcacc	tgctgcgctg	4260
gggcttcacc	acccccgaca	agaagcacca	gaaggagccc	cccttccctgt	ggatgggcta	4320
cgagctgcac	cccgacaagt	ggaccgtgca	gcccacgtg	ctgcccgaga	aggacagctg	4380
gaccgtgaac	gacatccaga	agctggtggg	caagctgaac	tgggcccagc	agatctacgc	4440
cggcatcaag	gtgcgccagc	tgtgcaagct	gctgcgcggc	accaaggccc	tgaccgaggt	4500
ggtgcccctg	accgaggagg	ccgagctgga	gctggccgag	aaccgcgaga	tcctgaagga	4560
gcccgtgcac	ggcgtgtact	acgaccccag	caaggacctg	atcgccgaga	tccagaagca	4620
gggcccaggc	cagtggacct	accagatcta	ccaggagccc	ttcaagaacc	tgaagaccgg	4680
caagtacgcc	cgcataaggg	gcgcccacac	caacgacgtg	aagcagctga	ccgaggccgt	4740
gcagaagatc	gccaccgaga	gcacgtgat	ctggggcaag	accccccaag	tcaagctgcc	4800
catccagaag	gagacctggg	aggcctggtg	gaccgagtac	tggcaggcca	cctggatccc	4860
cgagtgggag	ttcgtgaaca	ccccccccct	ggtgaagctg	tggtagcagc	tggagaagga	4920
gcccacatc	ggcgccgaga	ccttctacgt	ggacggcgcc	gccaaccgcg	agaccaagct	4980
gggcaaggcc	ggctacgtga	ccgaccgcgg	ccgcccagaag	gtggtgcccc	tgaccgacac	5040
caccaaccag	aagaccgagc	tgaggccat	ccacctggcc	ctgcaggaca	gcggcctgga	5100
ggtgaacatc	gtgaccgaca	gccagtacgc	cctgggcatc	atccaggccc	agcccgacaa	5160
gagcgagagc	gagctggtga	gccagatcat	cgagcagctg	atcaagaagg	agaagggtga	5220
cctggcctgg	gtgcccgcgc	acaagggcac	cgcgggcaac	gagcaggtgg	acggcctggt	5280
gagcgccggc	atccgcaagg	tgtgttccct	ggacggcatc	gacaaggccc	aggaggagca	5340
cgagaagtac	cacagcaact	ggcgcgccat	ggccagcgac	ttcaacctgc	cccccggtgt	5400
ggccaaggag	atcgtggcca	gctgcgacaa	gtgccagctg	aagggcgagg	ccatgcacgg	5460
ccaggtggac	tgacgccccg	gcacctggca	gctggcatgc	accacactgg	agggcaaggt	5520
gacctggtg	gccgtgcacg	tggccagcgg	ctacatcgag	gccgaggtga	tccccgccga	5580
gaccggccag	gagaccgcct	acttctgtgt	gaagctggcc	ggccgctggc	ccgtaaagac	5640
cgtgcacacc	gacaacggca	gcaacttcac	cagcaccacc	gtgaaggccg	cctgctggtg	5700
ggccggcatc	aagcaggagt	tcggcatccc	ctacaacccc	cagagccagg	gcgtgatcga	5760
gagcatgaac	aaggagctga	agaagatcat	cggccagggtg	cgcgaccagg	ccgagcacct	5820
gaagaccgcc	gtgcagatgg	ccgtgttcat	ccacaacttc	aagcgcaagg	gcggcatcgg	5880
cggctacagc	gccggcgagc	gcacgtggga	catcatcgcc	accgacatcc	agaccaagga	5940
gctgcagaag	cagatcacca	agatccagaa	cttccgcgtg	tactaccgcg	acagccgcga	6000
ccccgtgtgg	aagggccccg	ccaagctgct	gtggaagggc	gagggcgccg	tggatgacca	6060
ggacaacagc	gacatcaagg	tgggtgcccc	ccgcaaggcc	aagatcatcc	gcgactacgg	6120
caagcagatg	gccggcgagc	actgcgtggc	cagccgccag	gacgaggaca	tgggtaataa	6180
atggagcaag	tgtagcactg	tgggcccgcg	tgccatcagg	gagagaatgc	gccgcgcccc	6240
cgcagccgaa	ggagtgggtc	cagcttcaca	ggatagtgac	aaatacggcg	ctctgacatc	6300

cagcagtacc	cccgcataaca	acgccgactg	tgcattggctc	gaagcccagc	aggaagaaga	6360
ggagggtcga	ttccctgtga	ggcctcaggt	tcctctcaga	cctatgactt	acaaggccgt	6420
ggtcgacttg	tcttttttcc	ttgaagagaa	gggaggcttg	gaggggctga	tttacagcaa	6480
gaagcgccag	gatatttttg	atltgtgggt	ctacaacaca	caggggttact	tccccgattg	6540
gcagaactac	acccccggcc	caggcgtgcg	gtttcctctt	acctttgggt	ggtgctttta	6600
gctcgtgcct	gtcgatccta	gagaagtggg	ggaagcaaac	acaggcgaga	acaacagtct	6660
gctgcatact	atgagccttc	atggcatgga	ggattctcat	cgggaagtgc	ttcagtggaa	6720
attcgatagc	ttgctcgcta	gaagacacat	ggctcgggag	ttgcatactg	agtattataa	6780
ggattgttga	acacgtggcg	aattcttgct	gtgccttcta	gttgccagcc	atctgttggt	6840
tgcctctccc	ccgtgccttc	cttgaccctg	gaagggtgcca	ctcccactgt	cctttcctaa	6900
taaaatgagg	aaattgcata	gcattgtctg	agtaggtgtc	attctattct	gggggggtggg	6960
gtggggcagc	acagcaaggg	ggaggattgg	gaagacaata	gcaggcatgc	tggggatgag	7020
gtgggctcta	tgggtaccca	ggtgctgaag	aattgaccgc	gttcctcctg	ggccagaaag	7080
aagcaggcac	atccccctct	ctgtgacaca	ccctgtccac	gccccgtggt	cttagttcca	7140
gccccactca	taggacactc	atagctcagg	agggctccgc	cttcaatccc	acccgctaaa	7200
gtacttgag	cggctctctc	ctccctcctc	agccaccaa	accaaacctc	gcctccaaga	7260
gtgggaagaa	attaaagcaa	gataggctat	taagtgcaga	gggagagaaa	atgcctccaa	7320
catgtgagga	agtaatgaga	gaaatcatag	aatttcttcc	gttcctcgc	tcactgactc	7380
gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	7440
gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	7500
ggccaggaac	cgtaaaaagg	cgcgcttgct	ggcgtttttc	cataggctcc	gccccctga	7560
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	7620
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	7680
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	7740
ctgtagggtat	ctcagttcgg	tgtaggctcg	tcgctccaag	ctgggctgtg	tgcacgaacc	7800
ccccgttcag	cccgaccgct	gcgcttcttc	cggtaactat	cgtcttgagt	ccaaccgggt	7860
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	7920
tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	7980
agtattttgt	atctgcgctc	tgtgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	8040
ttgatccggc	aaacaaacca	cgcgtggtag	cgggtggtttt	tttgtttgca	agcagcagat	8100
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	8160
tcagtggaa	gaaaactcac	gttaagggat	tttgggtcatg	agattatcaa	aaaggatctt	8220
cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	8280
aacttggtct	gacagttacc	aatgcttaac	cagtgaggca	cctatctcag	cgatctgtct	8340
atttcggtca	tccatagttg	cctgactccg	gggggggggg	gcgctgagggt	ctgcctcgtg	8400
aagaagggtgt	tgtgactca	taccaggcct	gaatcgcccc	atcatccagc	cagaaagtga	8460
gggagccacg	gttgatgaga	gctttgttgt	aggtggacca	gttggtgatt	ttgaactttt	8520
gctttgccac	ggaacgggtct	gcggtgtcgg	gaagatgcgt	gatctgatcc	ttcaactcag	8580
caaaagttcg	atltattcaa	caaagccgcc	gtcccgtcaa	gtcagcgtaa	tgtctgcca	8640
gtgttacaac	caattaacca	atltctgatta	gaaaaactca	tcgagcatca	aatgaaactg	8700
caatttatct	atatcaggat	tatcaatacc	atatttttga	aaaagccgtt	tctgtaatga	8760
aggagaaaac	tcaccgaggc	agttccatag	gatggcaaga	tcctgggtatc	ggtctgcgat	8820
tccgactcgt	ccaacatcaa	tacaacctat	taatttcccc	tcgtcaaaaa	taagggttatc	8880
aagtgagaaa	tcaccatgag	tgacgactga	atccggtgag	aatggcaaaa	gcttatgcat	8940
ttctttccag	acttgttcaa	caggccagcc	attacgctcg	tcatacaaat	cactcgcactc	9000
aaccaaaccg	ttattcattc	gtgattgcgc	ctgagcgaga	cgaaatacgc	gatcgtgtgt	9060
aaaaggacaa	ttacaaacag	gaatcgaatg	caaccggcgc	aggaacactg	ccagcgcactc	9120
aacaatatct	tcacctgaat	caggatatct	ttctaatacc	tggaaatgctg	ttttcccggtg	9180
gacgcagtg	gtgagtaacc	atgcatactc	aggagtacgg	ataaaatgct	tgatgggtcgg	9240
aagaggcata	aattccgtca	gccagtttag	tctgaccatc	tcactctgta	catcattggc	9300
aacgctacct	ttgccatgtt	tcagaaacaa	ctctggcgca	tcgggcttcc	catacaatcg	9360
atagattgtc	gcacctgatt	gcccagacatt	atcgcgagcc	catttatacc	catataaatc	9420
agcatccatg	ttggaattta	atcgcgccct	cgagcaagac	gtttcccgtt	gaatatggct	9480
cataacaccc	cttgtattac	tgtttatgta	agcagacagt	tttatgttct	atgatgatat	9540
atltttatct	tgtgcaatgt	aacatcagag	atlttgagac	acaacgtggc	ttttcccccc	9600

```

ccccattat tgaagcattt atcaggggtta ttgtctcatg agcggatata tatttgaatg 9660
tatttagaaa aataaacaata taggggttcc gcgcacattt ccccgaaaag tgccacctga 9720
cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc 9780
ctttcgtc                                     9788

```

&lt;210&gt; 175

&lt;211&gt; 9407

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 175

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg ctttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
ccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgcag taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggcttg gggcctatac acccccgctt ccttatgcta 1140
taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggtat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggtctt tctccggtag cggcggagct tccacatccg agccctggc ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctectaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgacgc gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860
tgacgtcacc gtcgtcgaca tggcccgccg cagcgtgctg agcggcgggc agctggaccg 1920
ctgggagaag atccgcctgc gccccggcgg caagaagaag tacaagctga agcacatcgt 1980
gtgggccagc cgcgagctgg agcgttctgc cgtgaacccc ggcctgctgg agaccagcga 2040
gggctgccgc cagatcctgg gccagctgca gccagcctg cagaccggca gcgaggagct 2100
gcgcagcctg tacaacaccg tggccaccct gtactgcgtg caccagcgca tcgagatcaa 2160
ggacaccaag gaggccctgg acaagatcga ggaggagcag aacaagagca agaagaagge 2220
ccagcaggcc gccgccgaca ccggccacag caaccagggt agccagaact accccatcgt 2280
gcagaacatc cagggccaga tgggtcacca ggccatcagc ccccgacccc tgaacgcctg 2340

```



ggtgaaggtg	gtggaggaga	aggccttcag	ccccgaggtg	atccccatgt	tcagcgccct	2400
gagcgagggc	gccaccccc	aggacctgaa	caccatgctg	aacaccgtgg	gcggccacca	2460
ggccgccatg	cagatgctga	aggagaccat	caacgaggag	gccgccgagt	gggaccgcgt	2520
gcaccccggtg	cacgccggcc	ccatcgcccc	eggccagatg	cgcgagcccc	gcggcagcga	2580
catcgccggc	accaccagca	ccctgcagga	gcagatcggc	tggatgacca	acaaccccc	2640
catccccgtg	ggcgagatct	acaagcgctg	gatcatcctg	ggcctgaaca	agatcgtgcg	2700
catgtacagc	cccaccagca	tcctggacat	ccgccagggc	cccaaggagc	ccttcgcgca	2760
ctacgtggac	cgcttctaca	agaccctgcg	cgccgagcag	gccagccagg	aggtgaagaa	2820
ctggatgacc	gagaccctgc	tgggtgcagaa	cgccaacccc	gactgcaaga	ccatcctgaa	2880
ggccctgggc	cccgccgcca	ccctggagga	gatgatgacc	gcctgccagg	gcgtgggcgg	2940
ccccggccac	aaggcccgcg	tgctggccga	ggccatgagc	caggtgacca	acagcgccac	3000
catcatgatg	cagcgcgga	acttcgcgaa	ccagcgcaag	atcgtgaagt	gcttcaactg	3060
cggcaaggag	ggccacaccg	cccgcaactg	ccgcgcccc	cgcaagaagg	gctgctggaa	3120
gtgcggcaag	gagggccacc	agatgaagga	ctgcaccgag	cgacaggcta	ataacgccc	3180
ctgcgcctgg	ctggaggccc	aggaggagga	ggagggtggc	ttccccgtga	ccccccaggt	3240
gcccctgcgc	cccatgacct	acaaggccgc	cgtggacctg	agccacttcc	tgaaggagaa	3300
gggcggcctg	gagggcctga	tccacagcca	gcgccgccag	gacatcctgg	acctgtggat	3360
ctaccacacc	cagggctact	tccccgactg	gcagaactac	acccccggcc	ccggcgtgcg	3420
ctacccccctg	accttcggct	ggtgctacaa	gctggtgccc	gtggagcccc	acaaggtgga	3480
ggaggccaac	aagggcgaga	acaccagcct	gctgcacccc	gtgagcctgc	acggcatgga	3540
cgaccccgag	cgcgagggtg	tggagtggcg	cttcgacagc	cgcttggcct	tccaccacgt	3600
ggcccgcgag	ctgcaccccg	agtacttcaa	gaactgcgcc	cttctagaca	ccggcgccga	3660
cgacaccgtg	ctggaggaga	tgaacctgcc	cgcccgctgg	aagcccaaga	tgatcggcgg	3720
catcggcggc	ttcatcaagg	tgggcccagta	cgaccagatc	ctgatcgaga	tctgcggcca	3780
caaggccatc	ggcaccgtgc	tgggtggccc	cacccccgtg	aacatcatcg	gccgcaacct	3840
gctgaccag	atcggtgca	ccctgaactt	ccccatcagc	ccatcgaga	ccgtgcccgt	3900
gaagctgaag	cccggcatgg	acggccccaa	ggtgaagcag	tggcccccta	ccgaggagaa	3960
gatcaaggcc	ctgggtggaga	tctgcaccga	gatggagaag	gagggcaaga	tcagcaagat	4020
cggccccgag	aacccttaca	acacccccgt	gttcgccatc	aagaagaagg	acagcaccaa	4080
gtggcgcaag	ctgggtggact	tccgcgagct	gaacaagcgc	acccaggact	tctgggaggt	4140
gcagctgggc	atccccacc	ccgcggccct	gaagcagaag	aagagcgtga	ccgtgctgga	4200
cgtgggcgac	gcctacttca	gcgtgcccct	ggacaaggac	ttccgcaagt	acaccgcctt	4260
caccatcccc	agcatcaaca	acgagacccc	cggcacccgc	taccagtaca	acgtgctgcc	4320
ccagggtctg	aagggcagcc	ccgccatctt	ccagtgcagc	atgaccaaga	tcctggagcc	4380
cttccgcaag	cagaaccccg	acatcgtgat	ctaccagtac	atggaccacc	tgtacgtggg	4440
cagcgacctg	gagatcggcc	agcaccgcac	caagatcgag	gagctgcgcc	agcacctgct	4500
gcgctggggc	ttcaccaccc	ccgacaagaa	gcaccagaag	gagccccctt	tcctgtggat	4560
gggctacgag	ctgcaccccg	acaagtggac	cgtgcagccc	atcgtgctgc	ccgagaagga	4620
cagctggacc	gtgaacgaca	tccagaagct	ggtgggcaag	ctgaactggg	ccagccagat	4680
ctacgccggc	atcaagggtg	gccagctgtg	caagctgctg	cgcggcacca	aggccctgac	4740
cgaggtggtg	cccctgaccg	aggaggccga	gctggagctg	gccgagaacc	gcgagatcct	4800
gaaggagccc	gtgcacggcg	tgtactacga	ccccagcaag	gacctgatcg	ccgagatcca	4860
gaagcagggc	cagggccagt	ggacctacca	gatctaccag	gagcccttca	agaacctgaa	4920
gaccggcaag	tacgcccgca	tgaaggcgcg	ccacaccaac	gacgtgaagc	agctgaccga	4980
ggccgtgcag	aagatcgcca	ccgagagcat	cgtgatcttg	ggcaagaccc	ccaagttcaa	5040
gctgcccata	cagaaggaga	cctgggaggc	ctgggtggacc	gagtactggc	aggccacctg	5100
gatccccgag	tgggagtctg	tgaacacccc	ccccctggtg	aagctgtggt	accagctgga	5160
gaaggagccc	atcatcggcg	ccgagacctt	ctacgtggac	ggcgccgcca	accgcgagac	5220
caagctgggc	aaggccggct	acgtgaccga	ccgcggccgc	cagaagggtg	tgccccctgac	5280
cgacaccacc	aaccagaaga	ccgagctgca	ggccatccac	ctggccctgc	aggacagcgg	5340
cctggagggtg	aacatcgtga	ccgacagcca	gtacgcctg	ggcatcatcc	aggcccagcc	5400
cgacaagagc	gagagcgagc	tgggtgagcca	gatcatcgag	cagctgatca	agaaggagaa	5460
ggtgtacctg	gcctgggtgc	ccgcccacaa	gggcatcggc	ggcaacgagc	aggtggacgg	5520
cctggtgagc	gccggcatcc	gcaagggtgct	gttcctggac	ggcatcgaca	aggcccagga	5580
ggagcacgag	aagtaccaca	gcaactggcg	cgccatggcc	agcgacttca	acctgcccc	5640

cgtggtggcc aaggagatcg tggccagctg cgacaagtgc cagctgaagg gcgaggccat 5700  
gcacggccag gtggactgca gccccggcat ctggcagctg gcatgcaccc acctggaggg 5760  
caaggtgatc ctggtggccg tgcacgtggc cagcggctac atcgaggccg aggtgatccc 5820  
cgccgagacc ggccaggaga ccgcctactt cctgctgaag ctggccggcc gctggcccgt 5880  
gaagaccgtg cacaccgaca acggcagcaa cttcaccagc accaccgtga aggccgcctg 5940  
ctggtgggccc ggcatacaagc aggagtccg catccccctac aacccccaga gccagggcgt 6000  
gatcgagagc atgaacaagg agctgaagaa gatcatcggc caggtgcgcg accagggcca 6060  
gcacctgaag accgccgtgc agatggccgt gttcatccac aacttcaagc gcaagggcgg 6120  
catcgggcgc tacagcgccg gcgagcgcat cgtggacatc atcgccaccg acatccagac 6180  
caaggagctg cagaagcaga tcaccaagat ccagaacttc cgcgtgtact accgcgacag 6240  
ccgcgacccc gtgtggaagg gccccgcaa gctgctgtgg aagggcgagg gcgccgtgg 6300  
gatccaggac aacagcgaca tcaaggtgg gccccgccg aaggccaaga tcatccgca 6360  
ctacggcaag cagatggccg gcgacgactg cgtggccagc cgccaggacg aggactagga 6420  
attcttgctg tgccttctag ttgccagcca tctgttgttt gccctcccc cgtgccttcc 6480  
ttgaccctgg aaggtgccac tcccactgtc ctttccctaat aaaatgagga aattgcatcg 6540  
cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagca cagcaagggg 6600  
gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat gggtagccag 6660  
gtgctgaaga attgaccggt ttctctctgg gccagaaaga agcaggcaca tccccttctc 6720  
tgtgacacac cctgtccacg cccctggttc ttagttccag cccactcat aggacactca 6780  
tagctcagga gggctccgcc ttcaatccca cccgctaaag tacttgagc ggtctctccc 6840  
tccctcatca gccaccaaa ccaaacctag cctccaagag tgggaagaaa tttaaagcaag 6900  
ataggctatt aagtgcagag ggagagaaaa tgcctccaac atgtgaggaa gtaatgagag 6960  
aaatcataga atttcttccg ctctctcgt cactgactcg ctgcgctcg tcttccggt 7020  
gcggcgagcg gtatcagctc actcaaaagg ggtaatacgg ttatccacag aatcagggga 7080  
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 7140  
cgcggtgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg 7200  
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 7260  
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 7320  
tctcccttcg ggaagcgtgg cgcttttctca atgctcacgc tgtaggatc tcagttcgg 7380  
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 7440  
cgcttatcc ggtaactatc gtcttgagtc caaccgggta agacacgact tatcgccact 7500  
ggcagcagcc actggtaaca ggattagcag agcgaggatb gtaggcgggtg ctacagagtb 7560  
cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggtg tctgcgctct 7620  
gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 7680  
cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7740  
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg 7800  
ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 7860  
aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttggtctg acagttacca 7920  
atgcttaatc agtgaggcac ctatctcagc gatctgtcta ttctggtcat ccatagttgc 7980  
ctgactccgg gggggggggg cgctgaggtc tgcctcgtga agaaggtgtt gctgactcat 8040  
accaggcctg aatcgcccca tcatccagcc agaaagttag ggagccacgg ttgatgagag 8100  
ctttgttgta ggtggaccag ttggtgattt tgaacttttg ctttgccacg gaacggctctg 8160  
cgttgtcggg aagatgcgtg atctgatcct tcaactcagc aaaagttcga tttattcaac 8220  
aaagccgccc tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaaccaa 8280  
ttctgattag aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt 8340  
atcaatacca tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact caccgaggca 8400  
gttccatagg atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat 8460  
acaacctatt aatttcccct cgtcaaaaat aaggttatca agtgagaaat caccatgagt 8520  
gacgactgaa tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac 8580  
aggccagcca ttacgctcgt catcaaaatc actcgcatca accaaaccgt tattcattcg 8640  
tgattgcgc tgagcgagac gaaatacgcg atcgctgtta aaaggacaat taaaaacagg 8700  
aatcgaatgc aaccggcgca ggaacactgc cagcgcacat acaatatctt cacctgaatc 8760  
aggatattct tctaatacct ggaatgctgt tttcccgggg atcgagtgag agaggcataa 8820  
tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgtcag 8880  
ccagtttagt ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt 8940



```

cagaaacaac tctggcgcat cgggcttccc atacaatcga tagattgtcg cacctgattg 9000
cccgacatta tcgcgagccc atttataccc atataaatca gcatccatgt tgggaatttaa 9060
tcgcggcctc gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact 9120
gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 9180
acatcagaga ttttgagaca caacgtggct tcccccccc cccattatt gaagcattta 9240
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 9300
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 9360
catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc 9407

```

&lt;210&gt; 176

&lt;211&gt; 9785

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; plasmid pVR1012x/s containing HIV genes

&lt;400&gt; 176

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420
cccgcctggc tgaccgccc aacgaccccc cccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaaag acgcccccta ttgacgtcaa 600
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttccctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggatth ccaagtctcc accccattga 780
cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctcg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgtt ccttatgcta 1140
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt ccccggtgcc 1380
cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440
catgggctct tctccggtag cggcggagct tccacatccg agccctgggt ccatgcctcc 1500
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860
tgcagtcacc gtcgtcgaca tgggcgcccc cgccagcgtg ctgagcggcg gcaagttcga 1920
cgcctgggag aagatccgcc tgcgccccgg cggcaagaag aagtaccgca tgaagcacct 1980
gatctgggac agccgcgagc tggaccgctt cgccctgaac ccagcctgc tggagaccgc 2040
cgacggctgc cagcagatca tggagaagct gcagcccgcc ctgaagaccg gcaccgagga 2100

```

gatcaagagc ctgttcaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgacgt 2160  
gaaggacacc aaggaggccc tggacaagat cgaggagatc aagaacaaga gcaagcagaa 2220  
gacccagcag gccgcgcgcg acaccggcaa cagcagcaac gtgagccaga actaccccat 2280  
cgtgcagaac gcccagggcc agatgatcca ccagagcctg agcccccgca ccctgaacgc 2340  
ctgggtgaag gtgatcgagg agaaggcctt cagccccgag gtgatcccca tgttcagcgc 2400  
cctgagcgag ggcgccaccc cccaggacct gaacatgatg ctgaacatcg tgggcggcca 2460  
ccaggccgcc atgcagatgc tgaaggacac catcaacgag gaggccgcgc agtgggaccg 2520  
cctgcacccc gtgcacgcgc gccccatccc ccccgccag atgcgcgagc cccgcggcag 2580  
cgacatcgcc ggcaccacca gacccccca ggagcagatc ggctggatga ccggcaacc 2640  
ccccatcccc gtgggcgaca tctacaagcg ctggatcatc ctgggcctga acaagatcgt 2700  
gcgcagtac agccccgtga gcatcctgga catcaagcag ggccccaagg agcccttcgc 2760  
cgactacgtg gaccgcttct tcaagacctt ccgcgcgcgag caggccaccc aggacgtgaa 2820  
gaactggatg accgacaccc tgctgggtgca gaacgccaac cccgactgca aggccatcct 2880  
gcgcgccttg ggcgccggcg ccacctgga ggagatgatg accgcctgcc agggcggtgg 2940  
cggccccggc cacaaggccc gcgtgctggc cgaggccatg agccaggtgc agcacaccaa 3000  
catcatgatg cagcgcggca acttcaaggg ccagaagcgc atcaagtgt tcaactgcgg 3060  
caaggagggc cacctggccc gcaactgccg cgcctgcgc aagaagggt gctggaagt 3120  
cggcaaggag ggccaccaga tgaaggactg caccgagcga caggctaata gggaagatct 3180  
ggccttcccc caagggaagg ccagggaatt ttcttcagag cagaccagag ccaacagccc 3240  
caccagaaga gagcttcagg tttggggaag agacaacaac tccctctcag aagcaggagc 3300  
cgatagacaa ggaactgtat ccttttagctt cctcagatc actctttggc agcgaccct 3360  
cgtcacaata aagatagggg gccagctgaa ggaggccctt ctagacaccg gcgccgacga 3420  
caccgtgctg gaggagatga acctgcccgg ccgctggaag cccaagatga tcggcggcat 3480  
cggcggcttc atcaagggtg gccagtacga ccagatcctg atcgagatct gcggccacaa 3540  
ggccatcggc accgtgctgg tgggccccac ccccgtaac atcatcgcc gcaacctgt 3600  
gacccagatc ggctgcaccc tgaacttccc catcagcccc atcgagaccg tgcccgta 3660  
gctgaagccc ggcatggagc gcccgaagg gaagcagtgg cccctgaccg aggagaagat 3720  
caaggccctg gtggagatct gcaccgagat ggagaaggag ggcaagatca gcaagatcg 3780  
ccccgagaac cctacaaca ccccggtgtt cgccatcaag aagaaggaca gcaccaagt 3840  
gcgcaagctg gtggacttcc gcgagctgaa caagcgcacc caggacttct gggagggtgca 3900  
gctgggcata cccacccccg ccggcctgaa gcagaagaag agcgtgaccg tgctggacgt 3960  
gggcgacgcc tacttcagcg tgcccctgga caaggacttc cgcaagtaca ccgccttcac 4020  
catccccagc atcaacaacg agacccccgg catccgctac cagtacaacg tgctgcccc 4080  
gggctggaag ggcagccccg ccatcttcca gtgcagcatg accaagatcc tggagccctt 4140  
ccgcaagcag aaccccgaca tcgtgatcta ccagtacatg gaccacctgt acgtgggcag 4200  
cgacctggag atcgccacgc accgcaccaa gatcgaggag ctgcgccagc acctgctgcg 4260  
ctggggcttc accacccccg acaagaagca ccagaaggag cccccccttc tgtggatggg 4320  
ctacgagctg caccgcgaca agtggaacct gcagccatc gtgctgcccg agaaggacag 4380  
ctggaccgtg aacgacatcc agaagctggg gggcaagctg aactgggcca gccagatcta 4440  
cgccggcatc aagggtgcgc agctgtgcaa gctgctgcgc ggcaccaagg ccctgaccga 4500  
ggtggtgccc ctgaccgagg aggcgcagct ggagctggcc gagaaccgcg agatcctgaa 4560  
ggagcccgtg cagggcgtgt actacgacce cagcaaggac ctgatcgccg agatccagaa 4620  
gcagggccag ggccagtgga cctaccagat ctaccaggag cccttcaaga acctgaagac 4680  
cggcaagtac gcccgcatga agggcgccca caccaacgac gtgaagcagc tgaccgaggc 4740  
cgtgcagaag atcgccaccg agagcatcgt gatctggggc aagaccccca agttcaagct 4800  
gcccattccag aaggagacct gggaggcctg gtggaccgag tactggcagg ccacctggat 4860  
ccccgagtgg gagttcgtga acaccccccc cctggtgaag ctgtggtacc agctggagaa 4920  
ggagcccatc atcggcgcgc agaccttcta cgtggacggc gccgccaacc gcgagaccaa 4980  
gctgggcaag gccggctacg tgaccgaccg cggccgccag aagggtggtgc ccctgaccga 5040  
caccaccaac cagaagaccg agctgcaggc catccacctg gccctgcagg acagcggcct 5100  
ggaggtgaac atcgtgaccg acagccagta cgcctgggc atcatccagg cccagcccga 5160  
caagagcgag agcgagctgg tgagccagat catcgagcag ctgatcaaga aggagaaggt 5220  
gtacctggcc tgggtgcccg cccacaaggg catcggcggc aacgagcagg tggacggcct 5280  
ggtgagcgcc ggcattcgca aggtgctgtt cctggacggc atcgacaagg cccaggagga 5340  
gcacgagaag taccacagca actggcgcgc catggccagc gacttcaacc tgcccccg 5400

ggtggccaag gagatcgtgg ccagctgcga caagtgccag ctgaagggcg aggccatgca 5460  
cggccagggtg gactgcagcc cggcatctg gcagctggca tgcaccacc tggagggcaa 5520  
ggtgatcctg gtggccgtgc acgtggccag cggctacatc gaggccgagg tgatccccgc 5580  
cgagaccggc caggagaccg cctacttcct gctgaagctg gccggccgct ggcccgtgaa 5640  
gaccgtgcac accgacaacg gcagcaactt caccagcacc accgtgaagg ccgcctgctg 5700  
gtgggcccggc atcaagcagg agttcggcat cccctacaac cccagagacc agggcgtgat 5760  
cgagagcatg aacaaggagc tgaagaagat catcggccag gtgcgcgacc agggccagca 5820  
cctgaagacc gccgtgcaga tggccgtgtt catccacaac ttcaagcgca agggcggcat 5880  
cggcggctac agcgcggcg agcgcacgt ggacatcctc gccaccgaca tccagaccaa 5940  
ggagctgcag aagcagatca ccaagatcca gaacttccgc gtgtactacc gcgacagccg 6000  
cgaccccggtg tggaggggccc ccgccaagct gctgtggaag ggcgagggcg ccgtggtgat 6060  
ccaggacaac agcgacatca aggtggtgcc ccgcccgaag gccaaagatca tccgcgacta 6120  
cggcaagcag atggccggcg acgactgcgt ggccagccgc caggacgagg acatgggcag 6180  
caagtggagc aagagcagca tcgtgggctg gagggagggt agggagaggc tgaggcagac 6240  
cctggccgccc gccagaggcg tgggcgcccgt gagccaggac ctggagaagt acggcgcctg 6300  
gaccagcagc aacgtgaacc accctagctg cgcttggtg gagggccagg agggaggagg 6360  
ggtgggcttc cctgtgaggc ctcacgtgcc tctgaggcct atgacctaca agggcgccct 6420  
ggacctgagc cacttcctga aggagaagg cgccctggac ggcctgatct acagcaggag 6480  
gaggcaggag atcctggacc tgtgggtgta caacaccag ggctacttcc ctgactggca 6540  
gaactacaca cccggccctg gcgtgaggta ccctctgacc ttcggctggt gcttcaagct 6600  
ggtgcctgtg gacctgacg aggtggagca ggccaacgag ggcgagaaca acagcctgct 6660  
gcacctatg tgccagcacg gcatggacga cgaggagaag gagaccctga ggtggagggt 6720  
cgacagcagg ctggccctga ggcacagggc ccaggagatg caccctgagt tctacaagga 6780  
ctgctgaaca cgtggcgaat tcttgcctgt ccttctagtt gccagccatc tgttgtttgc 6840  
ccctcccccg tgccttcctt gacctggaa ggtgccactc ccaactgtct ttcctaataa 6900  
aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtggggtg 6960  
gggcagcaca gcaaggggga ggattgggaa gacaatagca ggcattgctg ggatgcggtg 7020  
ggctctatgg gtaccaggt gctgaagaat tgaccgggtt cctcctgggc cagaaagaag 7080  
caggcacatc ccttctctg tgacacacc tgtccacgcc cctggttctt agttccagcc 7140  
ccactcatag gacactcata gctcaggagg gctccgctt caatcccacc cgctaaagta 7200  
cttggagcgg tctctccctc cctcatcagc ccaccaaac aaacctagcc tccaagagt 7260  
ggaagaaatt aaagcaagat aggtatttaa gtgcagagg agagaaaatg cctccaacat 7320  
gtgaggaggt aatgagagaa atcatagaat ttcttccgct tcctcgtctc ctgactcgt 7380  
gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacgggt 7440  
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaaggc agcaaaaagg 7500  
caggaaccgt aaaaaggccg cgttgcctggc gtttttccat aggctccgcc cccctgacga 7560  
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 7620  
ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac 7680  
cggatacctg tccgccttct tcccttcggg aagcgtggcg ctttctcaat gctcacgctg 7740  
taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 7800  
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtag 7860  
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt 7920  
aggcgggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt 7980  
atthtggatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 8040  
atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac 8100  
gcgcagaaaa aaaggatctc aagaagatcc ttgatcttt tctacggggt ctgacgctca 8160  
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 8220  
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 8280  
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 8340  
tcgttcaccc atagttgcct gactccgggg gggggggcg ctgagggtct cctcgtgaag 8400  
aagggtgtgc tgactcatc caggcctgaa tcgccccatc atccagccag aaagtgagg 8460  
agccacgggt gatgagagct ttgtttaggt tggaccagtt ggtgattttg aacttttgct 8520  
ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat ctgatccttc aactcagcaa 8580  
aagttcgatt tattcaacaa agccgcccgt ccgtcaagtc agcgtaatgc tctgccagt 8640  
ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg agcatcaaat gaaactgcaa 8700

tttattcata	tcaggattat	caataccata	tttttgaaaa	agccgtttct	gtaatgaagg	8760
agaaaaactca	ccgaggcagt	tccataggat	ggcaagatcc	tggtatcggt	ctgcgattcc	8820
gactcgtcca	acatcaatac	aacctattaa	tttcccctcg	tcaaaaataa	ggttatcaag	8880
tgagaaatca	ccatgagtga	cgactgaatc	cggtgagaat	ggcaaaagct	tatgcatttc	8940
tttccagact	tggtcaacag	gccagccatt	acgctcgtca	tcaaaatcac	tcgcatcaac	9000
caaaccgtta	ttcattcgtg	attgcgcttg	agcgagacga	aatacgcgat	cgctgttaaa	9060
aggacaatta	caaacaggaa	tcgaatgcaa	ccggcgcagg	aacactgcca	gcgcatcaac	9120
aatattttca	cctgaatcag	gatattcttc	taatacctgg	aatgctgttt	tcccggggat	9180
cgcagtgggtg	agtaaccatg	catcatcagg	agtacggata	aaatgcttga	tggtcggaag	9240
aggcataaat	tccgtcagcc	agtttagtct	gaccatctca	tctgtaacat	cattggcaac	9300
gctacctttg	ccatgtttca	gaaacaactc	tggcgcacgc	ggcttcccat	acaatcgata	9360
gattgtcgca	cctgattgcc	cgacattatc	gcgagcccat	ttatacccat	ataaatcagc	9420
atccatgttg	gaatttaatc	gcggcctcga	gcaagacggt	tcccgttgaa	tatggctcat	9480
aacacccctt	gtattactgt	ttatgtaagc	agacagtttt	attgttcatg	atgatataat	9540
tttatcttgt	gcaatgtaac	atcagagatt	ttgagacaca	acgtggcttt	ccccccccc	9600
ccattattga	agcatttatc	agggttattg	tctcatgagc	ggatacatat	ttgaatgtat	9660
ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	cgaaaagtgc	cacctgacgt	9720
ctaagaaacc	attattatca	tgacattaac	ctataaaaaat	aggcgtatca	cgaggccctt	9780
tcgtc						9785